OM of: US-09-805-550-4 to: N_Geneseq_032802:* out_format : pfs

Date: Aug 31, 2002 5:59 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/US09805550/runat_29082002_160723_29077/app_query.fasta_1.897
-DB=N_Geneseq_032802 -QFMT=fastap -SUFFIX=p2n.rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09805550_@CGN1_1_0 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-805-550-4
Query length: 368
Database: N_Geneseq_032802:*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 294.590000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation ..		
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAD01231	+	1910.00	2435.61	2.6e-127	17			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC37705	+	1169.50	1488.82	1.4e-74	15			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC47900	+	1115.00	1419.08	1.1e-70	15			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC48229	+	931.50	1185.01	1.2e-57	14			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC38717	+	931.50	1185.00	1.2e-57	14			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAD01230	+	900.50	1144.74	2.0e-55	15			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:ABI99464	+	634.00	805.46	1.6e-36	12			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:AAT78310	+	603.50	767.65	2.1e-34	10			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAF21744	+	603.00	762.59	4.0e-34	17			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:AAT77781	+	574.50	730.97	2.3e-32	10			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC44960	+	563.50	718.72	1.1e-31	85			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABL03535	+	487.50	616.49	5.4e-26	1			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS78241	+	385.00	489.34	6.6e-19	9			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABL16782	+	375.00	459.88	2.9e-17	6			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABL03534	-	346.50	426.67	2.0e-15	4			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC43990	+	314.00	400.92	5.5e-14	72			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAA31521	+	281.50	362.97	7.2e-12	48			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:ABA60331	+	225.00	289.19	9.3e-08	5			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:ABA28583	+	225.00	289.19	9.3e-08	5			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAK08612	+	225.00	289.19	9.3e-08	5			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAK34493	+	225.00	289.19	9.3e-08	5			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI16748	+	225.00	289.19	9.3e-08	5			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI40213	+	225.00	289.19	9.3e-08	5			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAF08278	+	213.50	270.70	9.9e-07	86			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:AAS61710	+	196.50	253.65	8.8e-06	51			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC42513	+	195.50	241.38	4.3e-05	17			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH16080	+	177.00	214.81	0.0013	2			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAA60604	+	176.00	217.61	0.0009	15			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF57479	+	176.00	213.38	0.0015	2			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAF81716	+	176.00	210.68	0.0022	3			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI93885	+	174.00	209.11	0.0027	2			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH17711	+	174.00	209.11	0.0027	2			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS57824	+	171.00	224.82	0.0004	3			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH48066	+	170.50	203.38	0.0056	3			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS68865	+	169.00	206.89	0.0035	1			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAH16377	+	169.00	206.83	0.0036	1			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS68867	+	169.00	204.56	0.0048	2			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAS41980	+	169.00	201.10	0.0075	3			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAS41981	+	169.00	201.10	0.0075	3			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:AAQ60195	+	165.00	215.68	0.0011	39			
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS80744	+	163.00	197.70	0.0115				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAD00391	+	161.00	203.21	0.0057				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAD00394	+	160.50	196.57	0.0133				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:AAT27310	+	155.00	183.50	0.0713				
/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:AAX84092	+	155.00	183.50	0.0713				

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAD01231

seq_documentation_block:
ID AAD01231 standard; cDNA; 1702 BP.
XX
AC AAD01231;
XX
DT 04-OCT-2000 (first entry)
XX
DE Maize Rad23 protein #2 encoding cDNA.
XX
KW Rad23; maize; ATCC No: PTA-531; recombinant expression cassette; wheat;
KW transgenic plant; soybean; sunflower; sorghum; canola; modulator; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 106..1222
FT /*tag= a
FT /product= "Maize Rad23 protein #2"
FT /note= "Contained in ATCC No: PTA-531"
XX
PN WO200031268-A1.
XX
PD 02-JUN-2000.
XX
PF 12-OCT-1999; 99WO-US24129.
XX
PR 23-NOV-1998; 98US-0109728.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Mahajan PB, Tagliani L;
XX
DR WPI; 2000-400078/34.
DR P-PSDB; AAY71459.
XX
PT Isolated nucleic acid encoding maize RAD23 protein is used to modulate
PT the levels of polypeptides in plant or in assays for identifying
PT compounds that bind to and/or increase/decrease enzymatic activity of
PT catalytically active polypeptides -
XX
PS Claim 1e; Page 76-78; 82pp; English.
XX
CC The present sequence is the cDNA encoding maize Rad23 protein #2. It is
CC isolated from a Zea mays cell line, B73 callus tissue regenerated five
CC days after transfer of the callus from medium containing auzin to a
CC medium devoid of exogenous auxin. This cDNA is deposited under the ATCC
CC No: PTA-531. Maize Rad23 DNA sequence operably linked to a promoter, can
CC be used to construct a recombinant expression cassette. This expression
CC cassette can be used to generate a dicot or monocot transgenic plant
CC e.g., maize, soybean, sunflower, sorghum, canola, wheat, etc,. It can
CC also be used to modulate the levels of Rad23 polypeptide expression in a
CC plant or in assays to identify compounds, that bind to and/or modulate
CC the enzymatic activity of catalytically active polypeptides.
XX
SQ Sequence 1702 BP; 490 A; 373 C; 432 G; 407 T; 0 other;

alignment_scores:
Quality: 1910.00 Length: 368
Ratio: 5.190 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-805-550-4 x AAD01231 ..

```
Align seg 1/1  to: AAD01231  from: 1  to: 1702

      1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    106 ATGAAGCTGACGGTGAAGACCCTCAAGGGAACGCACTTCGAGATCCGGGT 155

     17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    156 GCAGCCCAACGACACGATTATGGCTGTGAAGAAGAATATAGAAGAGATAC 205

     34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    206 AAGGGAAAGACAGCTATCCATGGGGCCAACAACTGCTGATTTTCAATGGA 255

     51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    256 AAGGTCTTGAAAGATGAAAGTACATTGGAAGAGAATAAAGTCAATGAGGA 305

     67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrG 84
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    306 TGGGTTTCTAGTTGTCATGCTTAGTAAGGGTAAAACATCTGGTTCAACTG 355

     84 lyThrSerSerSerGlnHisSerAsnThrProAlaThrArgGlnAlaPro 100
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    356 GAACTTCATCTTCCCAGCACTCAAACACTCCTGCAACAAGGCAGGCACCT 405

    101 ProLeuGluAlaProGlnGlnAlaProGlnProProValAlaProIleTh 117
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    406 CCTCTAGAGGCCCCACAACAAGCTCCTCAACCCCCGGTGGCACCAATTAC 455

    117 rThrSerGlnProGluGlyLeuProAlaGlnAlaProAsnThrHisAspA 134
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    456 AACTTCTCAGCCTGAAGGACTTCCTGCACAGGCACCTAACACACATGACA 505

    134 snAlaAlaSerAsnLeuLeuSerGlyArgAsnValAspThrIleIleAsn 150
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    506 ATGCGGCATCAAATCTTCTGTCTGGAAGGAATGTTGACACAATAATTAAC 555

    151 GlnLeuMetGluMetGlyGlyGlySerTrpAspLysAspLysValGlnAr 167
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    556 CAGCTAATGGAGATGGGTGGGGGCAGTTGGGACAAAGATAAAGTCCAAAG 605

    167 gAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluTyrLeuT 184
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    606 GGCTCTCCGTGCCGCTTACAACAACCCCGAACGTGCTGTTGAATACCTCT 655

    184 yrSerGlyIleProValThrAlaGluIleAlaValProIleGlyGlyGln 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    656 ACTCTGGTATTCCAGTAACAGCTGAAATTGCTGTTCCAATTGGTGGTCAA 705

    201 GlyAlaAsnThrThrAspArgAlaProThrGlyGluAlaGlyLeuSerGl 217
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    706 GGGGCAAACACAACTGATCGAGCTCCTACTGGGGAAGCTGGTCTCTCTGG 755

    217 yIleProAsnThrAlaProLeuAspLeuPheProGlnGlyAlaSerAsnA 234
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    756 GATTCCAAACACCGCTCCACTAGATCTTTTCCCGCAGGGGGCTTCCAATG 805

    234 laGlyGlyGlyAlaGlyGlyGlyProLeuAspPheLeuArgAsnAsnPro 250
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    806 CTGGAGGTGGTGCTGGTGGTGGACCACTTGATTTTCTTAGAAACAATCCA 855

    251 GlnPheGlnAlaValArgGluMetValHisThrAsnProGlnIleLeuGl 267
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    856 CAGTTTCAAGCAGTTAGGGAAATGGTCCATACAAATCCACAAATTTTGCA 905

    267 nProMetLeuValGluLeuSerLysGlnAsnProGlnIleLeuArgLeuI 284
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    906 GCCTATGCTCGTTGAGTTGAGCAAGCAGAATCCTCAAATTCTAAGGTTGA 955

    284 leGluGluAsnHisAspGluPheLeuGlnLeuLeuAsnGluProPheGlu 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    956 TTGAGGAGAATCATGATGAGTTTCTTCAGTTACTAAATGAGCCCTTTGAA 1005

    301 GlyGlyGluGlyAspPheLeuAspGlnProGluGluAspGluMetProHi 317
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1006 GGCGGAGAGGGGGATTTCTTAGACCAACCTGAGGAGGATGAAATGCCTCA 1055

    317 sAlaIleSerValThrProGluGluGlnGluAlaIleGlyArgLeuGluS 334
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1056 TGCCATTAGTGTTACACCAGAGGAGCAGGAGGCCATTGGACGGCTTGAGT 1105

    334 erMetGlyPheAspArgAlaArgValIleGluAlaPheLeuAlaCysAsp 350
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1106 CCATGGGGTTCGACAGAGCACGCGTTATTGAAGCATTCTTAGCCTGCGAT 1155

    351 ArgAsnGluGluLeuAlaAlaAsnTyrLeuLeuGluHisAlaGlyGluGl 367
        ||||||||||||||||||||||||||||||||||||||||||||||||||
   1156 AGGAACGAGGAGCTAGCAGCAAACTATCTCCTTGAGCATGCTGGTGAGGA 1205

    367 uAsp 368
        ||||
   1206 AGAT 1209

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC37705

seq_documentation_block:
ID   AAC37705 standard; DNA; 1551 BP.
XX
AC   AAC37705;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 18371.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
```

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
           Quality: 1169.50            Length:    384
             Ratio:  3.748              Gaps:      9
 Percent Similarity:  81.250   Percent Identity: 64.323

alignment_block:
 US-09-805-550-4 x AAC37705   ..

 Align seg 1/1  to: AAC37705  from: 1  to: 1551

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         ||||||||||||||||||||||||||||||:::|||||||||||||||||
     140 ATGAAGCTCACCGTTAAAACTCTCAAGGGCAGTCATTTTGAAATTAGGGT 189

      17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
         |   |||:::|||||||||||||||||||||||||||||||||||::   |
     190 TTTACCCTCCGATACGATAATGGCGGTGAAGAAGAATATAGAAGATTCAC 239

      34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
         ||||||||||:::||||||   ||||||||||||||||||   ||||||
     240 AAGGAAAAGACAATTATCCATGTGGGCAGCAATTGTTGATTCACAATGGA 289

      51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
         ||||||||||||||||||:::::|||   ||||||||||||:::|||::
     290 AAAGTTTTGAAAGATGAAACTAGCCTGGTGGAGAACAAAGTTACGGAGGA 339

      67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrG 84
         :|||||||||||||||||||||||||:::|||:::::|||||||:::|
     340 AGGATTTCTCGTCGTTATGCTTAGCAAGAGCAAATCTGGAGGTTCAGCTG 389

      84 lyThrSerSerSerGlnHisSerAsn..................... 92
         ||   :::|||   |||   |||:::
     390 GTCAGGCTTCGGTTCAGACTTCTTCTGTTTCACAGCCTGTTTCCGCTACC 439

      93 ThrProAlaThrArgGlnAlaProProLeuGluAlaProGlnGlnAlaPr 109
         |||   :::|||:::   |||   |||      :::      |||:::
     440 ACATCTTCGACTAAGCCTGCAGCTCCA......TCTACAACCCAGTCTTC 483

     109 oGlnProProValAlaProIleThrThrSerGlnProGluGlyLeuProA 126
             |||::::::||||||   :::::::::            ||||
     484 ACCTGTTCCTGCTTCACCTATTCCTGCTCAAGAACAA.........CCGG 524

     126 laGlnAlaProAsnThrHisAspAsnAlaAlaSerAsnLeuLeuSerGly 142
         ||         :::|||:::   :::|||||||||:::|||:::||||||
     525 CGGCACAGACTGACACCTATGGACAAGCTGCGTCAACATTAGTTAGTGGT 574

     143 ArgAsnValAspThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySe 159
           :::::::::   ::::::::::|||:::|||||||||||||||||||
     575 AGTAGTCTTGAGCAAATGGTCCAACAGATAATGGAAATGGGAGGTGGAAG 624

     159 rTrpAspLysAspLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnP 176
         ||||||||||:::   |||   ||||||||||||||||||||||||||||
     625 TTGGGACAAAGAAACAGTTACTCGTGCACTTCGTGCAGCTTATAACAACC 674

     176 roGluArgAlaValGluTyrLeuTyrSerGlyIleProValThrAlaGlu 192
         ||||||||||||||:::|||||||||||||||||||   |||||||||
     675 CTGAGAGAGCAGTGGATTATTTGTACTCTGGAATTCCTCAAACTGCTGAA 724

     193 IleAlaValPro..............IleGlyGlyGlnGlyAlaAsnTh 204
         :::|||||||||              |||:::|||:::||||||
     725 GTAGCTGTGCCAGTACCTGAGGCTCAGATTGCAGGATCCGGTGCA..... 769

     204 rThrAspArgAlaProThrGlyGluAlaGlyLeuSerGlyIleProAsnT 221
                 ||||||::::::   |||      ||||||   ||||||:
     770 .........GCCCCTGTTGCACCTGCA......TCTGGAGGACCTAACT 803

     221 hrAlaProLeuAspLeuProProGlnGlyAlaSerAsnAlaGlyGlyGly 237
         :::|||||||||||||||||||||||||||||||||||||||||||||||
     804 CATCTCCTTTGGATTTATTTCCTCAGGAAACAGTTGCTGCTGCTGGCTCT 853

     238 AlaGlyGlyGlyProLeuAspPheLeuArgAsnAsnProGlnPheGlnAl 254
         :::      |||   |||:::||||||||||||||   |||||||||
     854 GGAGATCTTGGGACACTTGAATTTCTTAGAAACAATGACCAGTTCCAACA 903

     254 aValArgGluMetValHisThrAsnProGlnIleLeuGlnProMetLeuV 271
          :::|||   |||||||||:::|||||||||||||||||||||||||||||
     904 ATTACGCACCATGGTCCATTCCAACCCTCAAATTCTTCAGCCCATGCTTC 953

     271 alGluLeuSerLysGlnAsnProGlnIleLeuArgLeuIleGluGluAsn 287
           ||||||:::|||||||||||||||:::|||||||||||||:::||||||
     954 AGGAACTTGGAAAGCAAAACCCTCAACTTCTGAGGCTAATTCAAGAAAAC 1003

     288 HisAspGluPheLeuGlnLeuLeuAsnGluProPheGluGly......Gl 302
         :::   |||||||||||||||:::|||||||||:::||||||      ||
    1004 CAGGCAGAGTTTCTTCAGTTAGTAAATGAGCCCTATGAGGGTTCTGACGG 1053

     302 yGluGlyAspPheLeuAspGlnProGluGluAspGluMetProHisAlaI 319
         ||||||||||:::::::||||||||||||:::   |||||||||||||||
    1054 GGAGGGGGACATGTTCGATCAACCCGAGCAA...GAAATGCCACATGCCA 1100

     319 leSerValThrProGluGluGlnGluAlaIleGlyArgLeuGluSerMet 335
         ||:::|||||||||   |||||||||||||   |||||||||:::|||
    1101 TTAACGTTACCCCAGCAGAACAAGAAGCGATTCAACGGCTCGAGGCAATG 1150

     336 GlyPheAspArgAlaArgValIleGluAlaPheLeuAlaCysAspArgAs 352
         ||||||||||||||   ||||||||||||||||||||||||||||||||
    1151 GGATTTGATAGAGCCTTAGTAATAGAAGCGTTCCTCGCATGTGACCGAAA 1200

     352 nGluGluLeuAlaAlaAsnTyrLeuLeuGluHisAlaGlyGlu...GluA 368
         ||||||||||||||||||||||||||||||:::::|||:::   ||||
    1201 TGAAGAATTGGCAGCTAACTATCTATTGGAGAACTCAGGTGACTTTGAAG 1250

     368 sp 368
         ||
    1251 AC 1252

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC47900

seq_documentation_block:
ID   AAC47900 standard; DNA; 1549 BP.
XX
AC   AAC47900;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 55530.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
```

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

```
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
          Quality: 1115.00              Length:    374
            Ratio:  3.574                Gaps:        8
 Percent Similarity:  83.422   Percent Identity: 62.834

alignment_block:
 US-09-805-550-4 x AAC47900   ..

 Align seg 1/1  to: AAC47900  from: 1  to: 1549

      1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
        ||||||||||||||||||||||||||||||:::|||||||||||||||||
    232 ATGAAGCTCACTGTTAAGACTCTCAAGGGTAGCCATTTTGAGATTAGGGT 281

     17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
        |   |||:::|||||||||||||||||||||||||||||||||:::   |
    282 TCTTCCCACCGACACGATAATGGCGGTGAAGAAGAATATTGAAGATTCAC 331

     34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
        ||:::||||||:::||||||   ||||||||||||||||||   ||||||
    332 AAAGCAAAGACAACTATCCTTGTGGGCAGCAATTACTGATTCACAATGGA 381

     51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
        ||||||||||||||||||:::||||||   ||||||||||||:::|||::
    382 AAGGTTTTGAAAGATGAAACTACCTTGGTGGAGAACAAGGTTACCGAGGA 431

     67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrG 84
        :|||||||||||||||||||||||||:::||||||:::::::|||:::|
    432 GGGTTTTCTTGTCGTGATGCTTAGCAAGAGCAAAACTGCAAGTTCAGCTG 481

     84 lyThrSerSerSerGln...HisSerAsnThrProAlaThrArgGlnAla 99
        ||   ||||||:::|||      |||:::|||   :::|||   ::::::
    482 GTCCCTCTTCTACTCAGCCTACTTCTACCACGACATCTACCATATCTTCA 531

    100 ProProLeuGluAlaProGlnGlnAlaProGlnProProValAlaProIl 116
            |||   ||||||   ::::::   |||      :::::::||
    532 ACCACGCTTGCAGCTCCG...TCGACAACCCAGTCTATTGCTGTGCCGGC 578

    116 eThrThrSerGlnPro...GluGlyLeuProAlaGlnAlaProAsnThrH 132
         ::::::|||   |||   :::      |||:::          :::|||:
    579 TTCAAATTCTACTCCCGTTCAAGAACAACCAACGGCACAAAGTGACACCT 628

    132 isAspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnValAspThrIle 148
        ::   :::|||||||||:::|||:::||||||   :::::::::   :::
    629 ATGGTCAAGCTGCTTCAACTTTAGTTAGTGGCAGTAGTATTGAGCAAATG 678

    149 IleAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLysAspLysVa 165
        ::::::|||:::|||||||||||||||||||||||||||||:::   ||
    679 GTTCAACAAATAATGGAAATGGGAGGAGGCAGTTGGGACAAAGAAACGGT 728

    165 lGlnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluT 182
        |   ||||||||||||||||||||||||||||||||||||||||||:::|
    729 TACTCGTGCACTTCGTGCAGCATATAACAACCCTGAGAGAGCAGTGGATT 778

    182 yrLeuTyrSerGlyIleProValThrAlaGluIle...AlaValProIle 197
        ||||||||||||||||||   |||:::   |||   |||:::   :::
    779 ATCTATATTCTGGAATTCCTGAAACAGTAACCATTCCAGCAACTAATTTA 828

    198 GlyGlyGlnGlyAlaAsnThrThrAspArgAlaProThrGlyGluAlaGl 214
        :::|||   |||::::::          ||||||
    829 TCTGGAGTAGGATCTGGTAGAGAACTTACTGCTCCTCCTCCC........ 870

    214 yLeuSerGlyIleProAsnThrAlaProLeuAspLeuPheProGlnGlyA 231
            ||||||   ||||||::::::|||||||||||||||||||   |
    871 ....TCTGGAGGCCCTAATTCATCTCCTCTGGATTTGTTTCCCCAGGAAG 916

    231 laSerAsnAlaGlyGlyGlyAlaGlyGlyGlyProLeuAspPheLeuArg 247
        ||   :::   ::::::|||:::      |||   |||:::|||||||||
    917 CAGTTTCTGATGCAGCAGGTGGAGATCTTGGAACGCTTGAATTCCTCAGA 966

    248 AsnAsnProGlnPheGlnAlaValArgGluMetValHisThrAsnProGl 264
        :::|||   |||||||||   :::|||:::||||||::::::|||||||||
    967 GGCAATGATCAGTTCCAACAATTACGCTCCATGGTCAATTCCAACCCCCA 1016

    264 nIleLeuGlnProMetLeuValGluLeuSerLysGlnAsnProGlnIleL 281
        ||||||||||||||||||   ||||||:::|||||||||||||||||:::|
   1017 GATTCTGCAGCCTATGCTTCAAGAGCTCGGAAAGCAGAACCCCCAACTTC 1066

    281 euArgLeuIleGluGluAsnHisAspGluPheLeuGlnLeuLeuAsnGlu 297
        ||||||||||:::||||||:::   ||||||||||||||||||||||||
   1067 TGAGGCTAATTCAAGAGAACCAAGCCGAATTTCTTCAGTTACTAAACGAG 1116

    298 ProPheGluGlyGlyGluGly......AspPheLeuAspGlnProGluGl 312
        |||:::||||||::::::|||      |||::::::|||||||||:::::
   1117 CCCTACGAAGGATCTGACGGGGATGTGGATATCTTCGATCAACCTGATCA 1166

    312 uAspGluMetProHisAlaIleSerValThrProGluGluGlnGluAlaI 329
        :   ||||||||||||:::::::::|||||||||||||||||||||:::|
   1167 A...GAAATGCCCCACTCAGTCAACGTTACCCCTGAAGAGCAAGAATCAA 1213

    329 leGlyArgLeuGluSerMetGlyPheAspArgAlaArgValIleGluAla 345
        ||   |||||||||:::|||||||||||||||||   ||||||||||||
   1214 TTGAACGGCTTGAGGCAATGGGGTTTGATAGAGCAATAGTCATAGAAGCC 1263

    346 PheLeuAlaCysAspArgAsnGluGluLeuAlaAlaAsnTyrLeuLeuGl 362
        ||||||:::||||||||||||||||||||||||||||||||||||||||
   1264 TTCCTTTCCTGTGACCGTAACGAGGAATTGGCTGCAAACTATCTACTAGA 1313

    362 uHisAlaGlyGlu...GluAsp 368
        ||||:::::::::   ||||||
   1314 GCACTCAGCAGATTTTGAAGAC 1335

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC48229

seq_documentation_block:
ID   AAC48229 standard; DNA; 1424 BP.
XX
AC   AAC48229;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 56722.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
```

XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

```
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
            Quality:  931.50              Length:    380
              Ratio:  3.179                 Gaps:      8
 Percent Similarity:  77.105   Percent Identity: 52.632

alignment_block:
 US-09-805-550-4 x AAC48229   ..

 Align seg 1/1  to: AAC48229  from: 1  to: 1424

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         ||||||:::   |||||||||||||:::|||:::::|||||||||:::||
      79 ATGAAGATTTTCGTGAAGACTCTCAGTGGTTCGAACTTCGAGATCGAAGT 128

      17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
         |:::|||   |||   :::     ||||||     ||||||   ::::
     129 AAAACCTGCAGATAAGGTTTCTGATGTTAAAACGGCTATAGAAACTGTTA 178

      34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
         ::|||   :::   ||||||   :::::|||:::|||||||   :::|||
     179 AAGGTGCAGAA...TACCCTGCTGCTAAGCAGATGCTGATCCACCAAGGA 225

      51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
         ||||||||||||||||||:::||||||||||||||||:::|||   |||::
     226 AAGGTTCTAAAGGATGAGACCACATTGGAAGAGAACAATGTTGTTGAGAA 275

      67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSer.........G 81
         ::::|||:::|||:::||||||||||||   |||:::|||         :
     276 CAGTTTCATTGTTATCATGTTGTCCAAGACCAAGGCTTCTCCAAGTGGGG 325

      81 lySerThrGlyThrSerSerSerGlnHisSerAsnThrProAlaThrArg 97
         ::||||||:::::::::   :::      ::::::   |||   |||
     326 CATCAACCGCATCTGCACCAGCACCTAGTGCTACACAGCCACAGACAGTA 375

      98 GlnAlaProProLeuGluAlaPro...GlnGlnAlaProGlnProProVa 113

         :::|||   ::::::||||||      :::::::|||   |||
     376 GCTACACCTCAGGTTTCTGCTCCTACTGCCTCAGTTCCAGTTCCTACAAG 425

     113 lAlaProIleThrThrSerGlnProGluGlyLeuProAlaGlnAlaProA 130
         :::     |||::::::   |||        :::       :
     426 TGGAACTGCAACCGCTGCAGCTCCAGCTACTGCTGCTTCGGTTCAGACAG 475

     130 snThrHisAspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnValAsp 146
         ::::::::   :::||||||||||||||||:::::|||   :::::::::
     476 ATGTGTATGGTCAAGCAGCATCAAACCTTGTTGCTGGAACTACTTTAGAG 525

     147 ThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLysAs 163
         :::   :::::::|||:::::::::|||||||||||||||||||||:::||
     526 TCCACTGTTCAGCAAATTCTTGATATGGGTGGAGGTAGTTGGGACCGTGA 575

     163 pLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaV 180
         |   |||   |||||||||||||||||||:::|||||||||||||||||||
     576 CACTGTTGTTCGTGCTCTGAGAGCTGCGTTCAACAATCCTGAAAGAGCTG 625

     180 alGluTyrLeuTyrSerGlyIleProValThrAlaGluIle......... 193
         |||||||||||||||||||||||||:::   |||||||||
     626 TTGAATACCTGTACTCGGGTATCCCTGCTCAAGCTGAAATCCCACCAGTT 675

     194 ......AlaValProIleGlyGlyGlnGlyAlaAsn......ThrThrAs 206
               |||          |||   |||:::||||||      :::
     676 GCTCAAGCCCCAGCTACTGGGGAACAGGCAGCCAATCCTCTAGCACAGCC 725

     206 pArgAlaProThrGlyGluAlaGlyLeuSerGlyIleProAsnThrAlaP 223
         :::    ::::::   |||:::  :::|||   ||||||:::   |
     726 CCAACAAGCAGCAGCTCCAGCAGCTGCAACTGGTGGTCCAAACGCAAATC 775

     223 roLeuAspLeuPheProGlnGlyAlaSerAsnAlaGlyGlyGlyAlaGly 239
         |||||:::||||||||||||||||         |||   :::||||||||||
     776 CATTAAACCTGTTCCCCCAGGGCATGCCCGCTGCAGATGCTGGTGCTGGA 825

     240 GlyGlyProLeuAspPheLeuArgAsnAsnProGlnPheGlnAlaValAr 256
         :::|||   ||||||||||||||||||:::   ||||||||||||:::||
     826 GCTGGTAATCTTGATTTCCTGCGTAACAGTCAACAGTTCCAAGCCTTGCG 875

     256 gGluMetValHisThrAsnProGlnIleLeuGlnProMetLeuValGluL 273
         |   ||||||:::::::||||||||||||||||||||||||||||||  ||||
     876 AGCTATGGTACAAGCAAACCCACAAATTCTACAGCCTATGCTTCAGGAGC 925

     273 euSerLysGlnAsnProGlnIleLeuArgLeuIleGluGluAsnHisAsp 289
         ||:::|||||||||||||||:::::::|||||||||:::|||:::::::
     926 TCGGTAAACAAAACCCCCAGCTTGTGCGACTAATTCAAGAGCACCAGGCT 975

     290 GluPheLeuGlnLeuLeuAsnGluProPheGluGlyGlyGluGlyAspPh 306
         :::||||||:::|||:::|||||||||   |||   ||||||   :::
     976 GACTTCCTACGCTTGATAAACGAACCTGTCGAG...GGAGAAGAGAATGT 1022

     306 eLeuAspGlnProGluGluAspGluMetProHisAlaIleSerValThrP 323
         :::::::|||   |||         |||||||:::|||::::::|||||||
    1023 TATGGAACAGTTGGAAGCA...GCAATGCCACAAGCTGTTACCGTTACAC 1069

     323 roGluGluGlnGluAlaIleGlyArgLeuGluSerMetGlyPheAspArg 339
         ||||||||:::|||||||||   |||||||||:::|||||||||||||||
    1070 CTGAAGAGCGTGAAGCCATTGAACGGCTTGAAGGGATGGGGTTTGATCGT 1119

     340 AlaArgValIleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAl 356
         |||   |||:::|||:::|||:::||||||:::::|||||||||||||||
    1120 GCGATGGTCTTGGAAGTGTTCTTTGCGTGTAACAAGAATGAAGAACTTGC 1169

     356 aAlaAsnTyrLeuLeuGluHisAlaGlyGlu...GluAsp 368
         ||||||||||||||||:::|||      |||   ||||||
    1170 AGCTAATTACCTTCTAGATCACATGCATGAGTTTGAAGAT 1209

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC38717

seq_documentation_block:
ID   AAC38717 standard; DNA; 1426 BP.
XX
AC   AAC38717;
```

XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 21980.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

```
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
            Quality:  931.50              Length:    380
              Ratio:  3.179                Gaps:      8
 Percent Similarity:  77.105   Percent Identity: 52.632

alignment_block:
 US-09-805-550-4 x AAC38717   ..

 Align seg 1/1  to: AAC38717  from: 1  to: 1426

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         ||||||:::   ||||||||||||:::|||:::::|||||||||:::||
      79 ATGAAGATTTTCGTGAAGACTCTCAGTGGTTCGAACTTCGAGATCGAAGT 128

      17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
         |:::|||   |||   :::      ||||||      ||||||   ::::
     129 AAAACCTGCAGATAAGGTTTCTGATGTTAAAACGGCTATAGAAACTGTTA 178

      34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
         ::|||   :::   ||||||   :::::::|||:::|||||   :::|||
     179 AAGGTGCAGAA...TACCCTGCTGCTAAGCAGATGCTGATCCACCAAGGA 225

      51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
         ||||||||||||||||||:::|||||||||||||||||:::|||   |||::
     226 AAGGTTCTAAAGGATGAGACCACATTGGAAGAGAACAATGTTGTTGAGAA 275

      67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSer.........G 81
         ::::|||:::|||:::||||||||||   |||:::|||         :
     276 CAGTTTCATTGTTATCATGTTGTCCAAGACCAAGGCTTCTCCAAGTGGGG 325

      81 lySerThrGlyThrSerSerSerGlnHisSerAsnThrProAlaThrArg 97
         ::||||||:::::::::   :::      ::::::   |||   |||
     326 CATCAACCGCATCTGCACCAGCACCTAGTGCTACACAGCCACAGACAGTA 375

      98 GlnAlaProProLeuGluAlaPro...GlnGlnAlaProGlnProProVa 113
            :::|||   ::::::||||||      :::::::|||   |||
     376 GCTACACCTCAGGTTTCTGCTCCTACTGCCTCAGTTCCAGTTCCTACAAG 425

     113 lAlaProIleThrThrSerGlnProGluGlyLeuProAlaGlnAlaProA 130
          :::      |||::::::   |||            :::         :
     426 TGGAACTGCAACCGCTGCAGCTCCAGCTACTGCTGCTTCGGTTCAGACAG 475

     130 snThrHisAspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnValAsp 146
         ::::::::   :::||||||||||||||||:::::|||   :::::::::
     476 ATGTGTATGGTCAAGCAGCATCAAACCTTGTTGCTGGAACTACTTTAGAG 525

     147 ThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLysAs 163
         :::   ::::::|||:::::::::|||||||||||||||||||||:::||
     526 TCCACTGTTCAGCAAATTCTTGATATGGGTGGAGGTAGTTGGGACCGTGA 575

     163 pLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaV 180
         |   |||   ||||||||||||||||:::||||||||||||||||||||
     576 CACTGTTGTTCGTGCTCTGAGAGCTGCGTTCAACAATCCTGAAAGAGCTG 625

     180 alGluTyrLeuTyrSerGlyIleProValThrAlaGluIle......... 193
         ||||||||||||||||||||||||||:::   |||||||||
     626 TTGAATACCTGTACTCGGGTATCCCTGCTCAAGCTGAAATCCCACCAGTT 675

     194 ......AlaValProIleGlyGlyGlnGlyAlaAsn......ThrThrAs 206
               |||          |||   |||:::||||||      :::
     676 GCTCAAGCCCCAGCTACTGGGGAACAGGCAGCCAATCCTCTAGCACAGCC 725

     206 pArgAlaProThrGlyGluAlaGlyLeuSerGlyIleProAsnThrAlaP 223
          :::      ::::::   |||:::   :::|||   ||||||:::   |
     726 CCAACAAGCAGCAGCTCCAGCAGCTGCAACTGGTGGTCCAAACGCAAATC 775

     223 roLeuAspLeuPheProGlnGlyAlaSerAsnAlaGlyGlyGlyAlaGly 239
         |||||:::||||||||||||||||        |||   :::||||||||||
     776 CATTAAACCTGTTCCCCCAGGGCATGCCCGCTGCAGATGCTGGTGCTGGA 825

     240 GlyGlyProLeuAspPheLeuArgAsnAsnProGlnPheGlnAlaValAr 256
         :::|||   ||||||||||||||||||:::   ||||||||||||:::||
     826 GCTGGTAATCTTGATTTCCTGCGTAACAGTCAACAGTTCCAAGCCTTGCG 875

     256 gGluMetValHisThrAsnProGlnIleLeuGlnProMetLeuValGluL 273
         |   ||||||:::::::||||||||||||||||||||||||||   ||||
     876 AGCTATGGTACAAGCAAACCCACAAATTCTACAGCCTATGCTTCAGGAGC 925

     273 euSerLysGlnAsnProGlnIleLeuArgLeuIleGluGluAsnHisAsp 289
         ||:::|||||||||||||||:::::|||||||||:::|||::::::
     926 TCGGTAAACAAAACCCCCAGCTTGTGCGACTAATTCAAGAGCACCAGGCT 975

     290 GluPheLeuGlnLeuLeuAsnGluProPheGluGlyGlyGluGlyAspPh 306
         :::|||||||:::|||:::|||||||||   |||   ||||||   :::
     976 GACTTCCTACGCTTGATAAACGAACCTGTCGAG...GGAGAAGAGAATGT 1022

     306 eLeuAspGlnProGluGluAspGluMetProHisAlaIleSerValThrP 323
          :::::::|||   |||         ||||||:::|||:::::|||||||
    1023 TATGGAACAGTTGGAAGCA...GCAATGCCACAAGCTGTTACCGTTACAC 1069

     323 roGluGluGlnGluAlaIleGlyArgLeuGluSerMetGlyPheAspArg 339
         ||||||||:::|||||||||   |||||||||:::|||||||||||||||
    1070 CTGAAGAGCGTGAAGCCATTGAACGGCTTGAAGGGATGGGGTTTGATCGT 1119
```

```
     340 AlaArgValIleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAl 356
         |||   |||:::|||:::|||:::||||||::::::|||||||||||||
    1120 GCGATGGTCTTGGAAGTGTTCTTTGCGTGTAACAAGAATGAAGAACTTGC 1169

     356 aAlaAsnTyrLeuLeuGluHisAlaGlyGlu...GluAsp 368
         ||||||||||||||||:::|||      |||   ||||||
    1170 AGCTAATTACCTTCTAGATCACATGCATGAGTTTGAGGAT 1209

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAD01230

seq_documentation_block:
ID   AAD01230 standard; cDNA; 1522 BP.
XX
AC   AAD01230;
XX
DT   04-OCT-2000  (first entry)
XX
DE   Maize Rad23 protein #1 encoding cDNA.
XX
KW   Rad23; maize; ATCC No: PTA-530; recombinant expression cassette; wheat;
KW   transgenic plant; soybean; sunflower; sorghum; canola; modulator; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             58..1275
FT                   /*tag= a
FT                   /product= "Maize Rad23 protein #1"
FT                   /note= "Contained in ATCC No: PTA-530"
XX
PN   WO200031268-A1.
XX
PD   02-JUN-2000.
XX
PF   12-OCT-1999;   99WO-US24129.
XX
PR   23-NOV-1998;   98US-0109728.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Mahajan PB,  Tagliani L;
XX
DR   WPI; 2000-400078/34.
DR   P-PSDB; AAY71458.
XX
PT   Isolated nucleic acid encoding maize RAD23 protein is used to modulate
PT   the levels of polypeptides in plant or in assays for identifying
PT   compounds that bind to and/or increase/decrease enzymatic activity of
PT   catalytically active polypeptides -
XX
PS   Claim 1e; Page 73-75; 82pp; English.
XX
CC   The present sequence is the cDNA encoding maize Rad23 protein #1. It is
CC   isolated from V5 root tissue of a Zea mays cell line B73, infested with
CC   corn root worm. This cDNA is deposited under the ATCC No: PTA-530. Maize
CC   Rad23 DNA sequence operably linked to a promoter can be used to construct
CC   a recombinant expression cassette. This expression cassette can be used
CC   to generate a dicot or monocot transgenic plant e.g., maize, soybean,
CC   sunflower, sorghum, canola, wheat, etc,. It can also be used to modulate
CC   the levels of Rad23 polypeptide expression in a plant or in assays to
CC   identify compounds, that bind to and/or modulate the enzymatic activity
CC   of catalytically active polypeptides.
XX
SQ   Sequence 1522 BP; 430 A; 375 C; 370 G; 347 T; 0 other;


alignment_scores:
           Quality:  900.50            Length:    402
             Ratio:  3.063               Gaps:      7
 Percent Similarity:  73.134   Percent Identity: 49.005

alignment_block:
 US-09-805-550-4 x AAD01230   ..
```

```
Align seg 1/1  to: AAD01230  from: 1  to: 1522

      1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
        |||||||||:::|||||||||||||||||||||||:::|||||||||:::::
     58 ATGAAGCTTAACGTCAAGACCCTCAAGGGCACCAACTTCGAGATCGAGGC 107

     17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
        :::|||:::   ::::::      ||||||:::   ||||||      |
    108 GAGCCCCGATGCATCGGTTGCTGATGTGAAGAGGATCATTGAGACCACTC 157

     34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
        |||||:::::::::|||         ||||||:::||||||::::::|||
    158 AAGGTCAGAGTACCTACCGGGCGGACCAGCAAATGCTCATATACCAAGGG 207

     51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
        |||:::||||||||||||:::||||||||||:::|||   |||   |||::
    208 AAAATTCTCAAGGATGAAACCACTTTGGAAAGCAACGGAGTTGCTGAGAA 257

     67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrG 84
        ::::|||||||||:::|||||||||||||:::|||:::|||   |||:::|
    258 CAGCTTCCTTGTTATAATGTTGTCCAAGGCTAAGGCATCA...TCGAGTG 304

     84 lyThrSerSerSerGlnHisSerAsnThrProAlaThrArgGlnAlaPro 100
        ||:::|||::::::      :::::::::|||||||||          |||
    305 GAGCTTCTACCGCTACTACTGCAAAAGCTCCTGCAACTCTGGCCCAACCT 354

    101 ..............ProLeuGluAlaProGlnGlnAlaPro...GlnPr 111
                       |||      :::      :::::|||   |||
    355 GCTGCCCCTGTGGCCCCTGCTGCATCAGTTGCAAGAACACCAACACAGGC 404

    111 oProValAlaProIleThrThrSerGlnProGluGlyLeuProAlaGlnA 128
         |||||||||         |||:::   |||:::      |||      |
    405 TCCTGTTGCCACAGCTGAAACGGCACCTCCAAGTGTCCAACCTCAGGCTG 454

    128 laPro.............................AsnThrHisAspAsn 134
        |||||                           :::::::::::::::
    455 CTCCAGCTGCTACGGTTGCTGCTACTGATGATGCTGATGTGTACAGTCAG 504

    135 AlaAlaSerAsnLeuLeuSerGlyArgAsnValAspThrIleIleAsnGl 151
        ||||||||||||||||:::   |||:::|||::::::      |||:::||
    505 GCAGCTTCAAACCTTGTATTTGGCAACAATCTAGAACAGACTATCCAACA 554

    151 nLeuMetGluMetGlyGlyGlySerTrpAspLysAspLysValGlnArgA 168
        |::::::::::|||||||||||||:::|||:::::|||   |||   ||||
    555 AATTCTTGACATGGGTGGTGGTACATGGGAACGTGATACTGTTGTTCGTG 604

    168 laLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluTyrLeuTyr 184
        ||||||||||||||||||||||||||||||||||:::::|||||||||
    605 CTCTACGTGCTGCATACAATAACCCCGAGAGAGCTATAGACTACCTGTAT 654

    185 SerGlyIleProValThrAlaGlu............IleAlaValProIl 197
        ||||||||||||   ::::::|||                  :::|||
    655 TCTGGAATTCCTGAGAATGTGGAGGCTCAGCCTGTTGCCCGAGCACCTGC 704

    197 eGlyGlyGlnGlyAlaAsnThrThrAspArgAlaProThrGlyGluAla. 213
         :::||||||   :::|||            :::|||:::      |||
    705 TGCTGGCCAACAAACAAATCAGCAGGCCGCATCACCCGCTCAGCCAGCAG 754

    214 .....................GlyLeuSerGlyIleProAsnThrAla 222
                               |||:::   ||||||:::
    755 TTGCATTGCCAGTGCAGCCATCACCTGCCTCTGCAGGGCCTAATGCAAAT 804

    223 ProLeuAspLeuProProGlnGlyAlaSerAsnAlaGlyGly........ 236
        ||||||:::||||||||||||||||:::   ::::::|||:::
    805 CCTTTGAACCTTTTTCCTCAGGGTGTTCCAAGTGGTGGGTCCAACCCAGG 854

    237 ..........GlyAlaGlyGlyGlyProLeuAspPheLeuArgAsnAsnP 250
                  |||||||||:::|||   ||||||   ||||||:::   |
    855 TGTTGTTCCAGGTGCAGGATCTGGTGCTCTTGATGCCTTGCGACAGCTTC 904

    250 roGlnPheGlnAlaValArgGluMetValHisThrAsnProGlnIleLeu 266
        ||||||||||||||:::   ::::::|||::::::|||||||||||||||
```

```
      905 CACAGTTTCAAGCACTCCTTCAGTTAGTCCAGGCTAATCCTCAAATCTTG 954

      267 GlnProMetLeuValGluLeuSerLysGlnAsnProGlnIleLeuArgLe 283
      955 CAGCCAATGCTTCAAGAGCTAGGTAAACAAAACCCACAAATTCTGCGGTT 1004

      283 uIleGluGluAsnHisAspGluPheLeuGlnLeuLeuAsnGluProPheG 300
     1005 GATTCAGGAAAATCAAGCTGAGTTTCTCCGCTTGGTGAATGAATCTCCTG 1054

      300 luGlyGlyGluGlyAspPheLeuAspGlnProGluGluAspGluMetPro 316
     1055 AGGGTGGTCCTGGAGGGAACATACTAGGTCAACTGGCAGCTGCTGTGCCA 1104

      317 HisAlaIleSerValThrProGluGluGlnGluAlaIleGlyArgLeuGl 333
     1105 CAAACGCTGACAGTTACCCCAGAGGAACGGGAGGCTATCCAGCGGCTCGA 1154

      333 uSerMetGlyPheAspArgAlaArgValIleGluAlaPheLeuAlaCysA 350
     1155 GGGAATGGGGTTCAACCGTGAGCTTGTGCTAGAAGTTTTCTTTGCATGCA 1204

      350 spArgAsnGluGluLeuAlaAlaAsnTyrLeuLeuGluHisAlaGlyGlu 366
     1205 ACAAGGACGAAGAGCTTACAGCCAACTACCTCCTGGATCATGGCCATGAG 1254

      367 GluAsp 368
     1255 TTTGAC 1260
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:ABI99464

seq_documentation_block:

ID ABI99464 standard; cDNA; 1251 BP.
XX
AC ABI99464;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:414.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR P-PSDB; ABB57171.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 1129-1132; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 1251 BP; 361 A; 323 C; 305 G; 262 T; 0 other;

alignment_scores:
Quality: 634.00 Length: 436
Ratio: 2.314 Gaps: 11
Percent Similarity: 62.844 Percent Identity: 33.486

alignment_block:
US-09-805-550-4 x ABI99464 ..

Align seg 1/1 to: ABI99464 from: 1 to: 1251

```
        1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
        1 ATGCAGGTCACCCTGAAGACCCTTCAGCAGCAGACCTTCAAGATCGACAT 50

       17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
       51 CGACCCGGAGGAGACGGTAAAGGCATTGAAAGAGAAGATTGAATCTGAAA 100

       34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
      101 AGGGGAAAGATGCCTTTCCGGTAGCAGGTCAGAAGTTAATTTATGCCGGC 150

       51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
      151 AAAATCCTCAGTGATGATACTGCTCTCAAAGAATATAAAATTGATGAGAA 200

       67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrG 84
      201 AAACTTTGTGGTGGTTATGGTGACAAAACCCAAAGCAGTGACAACAGCAG 250

       84 lyThrSerSerSerGlnHisSerAsnThrProAla.....ThrArg 97
      251 TGCCAGCTACAACCCAGCCATCAAGTACTCCCAGCCCCACTACAGTCAGT 300

       98 GlnAlaProProLeuGluAlaProGln...GlnAlaProGlnProProVa 113
      301 TCTTCCCCAGCAGTGGCTGCGGCCCAGGCTCCAGCTCCCACCCCTGCTCT 350

      113 lAlaProIleThrThr..... 118
      351 GGCTCCCACTTCCACTCCTGCCTCCACTACTCCAGCCTCCACAACAGCCT 400

      119 .....SerGlnProGlu.....Gly 123
      401 CTTCTGAACCCGCACCTGCTGGTGCCACTCAGCCTGAGAAACCTGCAGAA 450

      124 LeuProAlaGlnAlaPro..... 129
      451 AAGCCAGCCCAGACACCAGTGCTTACTAGCCCAGCACCAGCTGACAGTAC 500

      130 .....AsnThrHisAspAsnAlaAlaSerAsnL 139
      501 ACCAGGAGATTCTTCCCGGTCAAATCTTTTTGAAGATGCAACAAGTGCCC 550

      139 euLeuSerGlyArgAsnValAspThrIleIleAsnGlnLeuMetGluMet 155
      551 TTGTGACAGGTCAGTCTTATGAGAATATGGTAACTGAGATCATGTCAATG 600

      156 GlyGlyGlySerTrpAspLysAspLysValGlnArgAlaLeuArgAlaAl 172
```

```
     601 GGC.........TATGAACGAGAACAAGTAATTGCAGCCCTGAGAGCCAG 641

     172 aTyrAsnAsnProGluArgAlaValGluTyrLeuTyrSerGlyIleProV 189
         ::::||||||||||:::||||||||||||||||||      |||||||||
     642 CTTCAACAACCCTGACAGAGCTGTGGAATATCTTCTAATGGGAATCCCTG 691

     189 alThrAlaGluIleAlaValProIleGlyGlyGlnGlyAlaAsnThrThr 205
               |||      :::   :::
     692 GAGATAGAGAAAGTCAGGCTGTGGTTGAC..................... 720

     206 AspArgAlaProThrGlyGluAlaGlyLeuSerGlyIleProAsnThrAl 222
               |||      :::|||      :::|||   |||::::::
     721 .........CCTCCTCCTCAGGCTGTGAGTACTGGAACTCCTCAGTCT.. 759

     222 aProLeuAspLeuPheProGlnGlyAlaSerAsnAlaGly.......... 235
                     |||      |||:::   |||:::
     760 ...............CCAGCAGTAGCTGCAGCTGCAGCAACCACGACAG 793

     236 ...........GlyGlyAlaGlyGlyGlyProLeuAspPheLeuArgAsn 248
                    :::||||||   ||||||:::||||||||||||||
     794 CAACTACAACAACCACTTCTGGAGGCCACCCCCTTGAATTTTTACGGAAT 843

     249 AsnProGlnPheGlnAlaValArgGluMetValHisThrAsnProGlnIl 265
         :::|||||||||||||   :::|||:::::::::::::   ||||||:::::
     844 CAGCCTCAGTTTCAACAGATGAGACAAATTATCCAGCAGAATCCTTCTTT 893

     265 eLeuGlnProMetLeuValGluLeuSerLysGlnAsnProGlnIleLeuA 282
         :|||      :::|||   ::::::::::::::||||||||||:::|||:
     894 GCTGCCAGCCTTGCTACAGCAGATAGGTCGGGAGAACCCTCAGCTGCTGC 943

     282 rgLeuIleGluGluAsnHisAspGluPheLeuGlnLeuLeuAsnGluPro 298
         ::   |||:::::::::::::::::::|||:::|||:::||||||||||||
     944 AGCAAATTAGCCAACACCAGGAGCATTTTATTCAGATGCTGAATGAACCA 993

     299 PheGlu.......................................... 300
           ::: .
     994 GTTCAGGAAGCAGGCGGTCAAGGCGGAGGAGGCGGTGGCGGAGGAGGCGG 1043

     301 ....GlyGlyGluGlyAspPheLeuAspGlnProGluGluAspGluMetP 316
             ||||||   |||      :::   :::      :::   :::|||
    1044 TGGCGGAGGAGGCGGCGGAGGAATTGCAGAAGCTGGAAGTGGGCACATGA 1093

     316 roHisAlaIleSerValThrProGluGluGlnGluAlaIleGlyArgLeu 332
           :::   |||:::|||||||||:::|||:::|||||||||   ||||||
    1094 ATTAC...ATTCAAGTAACACCTCAGGAGAAAGAAGCTATAGAACGGTTA 1140

     333 GluSerMetGlyPheAspArgAlaArgValIleGluAlaPheLeuAlaCy 349
         :::::::::||||||   ::::::   ||||||:::|||::::::|||||
    1141 AAGGCATTAGGATTTCCTGAAGGACTTGTGATTCAAGCATATTTGCTTG 1190

     349 sAspArgAsnGluGluLeuAlaAlaAsnTyrLeuLeuGluHisAlaGlyG 366
         |:::::|||||||:::||||||||||||:::||||||::::::      :
    1191 TGAGAAGAATGAGAATCTGGCTGCCAACTTTCTTCTACAGCAGAACTTTG 1240

     366 luGluAsp 368
         ::||||||
    1241 ATGAAGAC 1248

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:AAT78310

seq_documentation_block:
ID   AAT78310 standard; cDNA; 1092 BP.
XX
AC   AAT78310;
XX
DT   10-OCT-1997  (first entry)
XX
DE   E6AP-binding protein cln42 encoding cDNA.
XX
KW   Human; cell differentiation; survival; carcinoma; psoriasis;
KW   ubiquitination; p53; tumour suppressor; homeostasis; papilloma virus;
KW   epithelial cell; acne; icthyosis; aphthous ulcer; hair growth;
KW   antibody; cell proliferation; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1092
FT                   /*tag= a
FT                   /product= cln42
FT                   /note= "E6AP-binding protein"
XX
PN   WO9640767-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09040.
XX
PR   07-JUN-1995;   95US-0484878.
XX
PA   (MITO-) MITOTIX INC.
XX
PI   Beer-Romero PL,  Draetta G,  Rolfe M;
XX
DR   WPI; 1997-087053/08.
DR   P-PSDB; AAW23658.
XX
PT   E6AP-binding proteins and related nucleic acid - useful for
PT   modulating cell differentiation, survival etc., partic. for
PT   treatment and diagnosis of carcinoma, psoriasis, etc
XX
PS   Claim 12; Page 66-68; 83pp; English.
XX
CC   The present sequence encodes the E6AP-binding protein cln42. E6AP
CC   mediates ubiquitination and so the inactivation of e.g. p53 tumour
CC   suppressor. The E6AP-binding protein is likely to be involved in normal
CC   cell homeostasis and in the pathogenesis of proliferative and
CC   differentiation disorders, e.g. regulation of gene expression or the
CC   cell cycle, modification of cell surface receptors, biogenesis of
CC   ribosomes and DNA repair. The protein, which can optionally be generated
CC   in vivo by gene therapy, may useful in treatment and prevention of
CC   papilloma virus infected transformed cells and carcinoma, and may also
CC   be used to regulate epithelial cell processes more generally, e.g. in
CC   cases of psoriasis, acne, icthyosis, aphthous ulcers. It may also be
CC   used to control wound healing, and inhibit growth of hair. It can also
CC   be used to generate antibodies which are used in immunoassays to
CC   determine the protein levels. The encoding nucleic acid can be used to
CC   prepare recombinant proteins and oligonucleotides useful as probes and
CC   primers for diagnostic detection of (mutant) mRNA for the protein in
CC   (transformed) cells, also for antisense therapy. It can also be used
CC   for the detection of mutations in E6AP-binding protein-encoding genes,
CC   mis-expression of these genes or mis-incorporation of the protein in a
CC   transcription regulatory complex containing E6AP can be used to assess
CC   risk of disorders characterised by cell proliferation.
XX
SQ   Sequence 1092 BP; 260 A; 329 C; 324 G; 179 T; 0 other;


alignment_scores:
            Quality:  603.50              Length:    392
              Ratio:  2.269                 Gaps:      8
 Percent Similarity:  67.857   Percent Identity: 33.418

alignment_block:
 US-09-805-550-4 x AAT78310   ..

 Align seg 1/1  to: AAT78310  from: 1  to: 1092

       3 LeuThrValLysThrLeuLysGlyThrHisPheGluIleArgValGlnPr 19
         :::|||:::|||||||||:::      |||:::||||||:::::||
      13 ATCACGCTCAAAACGCTGCAGCAGCAGACCTTCAAGATCCGCATGGAGCC 62

      19 oAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleGlnGlyL 36
         |:::::|||:::   ::::::|||::::::|||||      :::|||:
      63 TGACGAGACGGTGAAGGTGCTAAAGGAGAAGATAGAAGCTGAGAAGGGTC 112
```

```
 36 ysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGlyLysVal 52
    ::|||:::::::|||   :::   |||   |||||||:::   ||||||::: 
113 GTGATGCCTTCCCCGTGGCTGGACAGAAACTCATCTATGCCGGCAAGATC 162

 53 LeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAspGlyPh 69
    |||:::|||:::      :::::::::   ::::::::::|||   :::|| 
163 TTGAGTGACGATGTCCCTATCAGGGACTATCGCATCGATGAGAAGAACTT 212

 69 eLeuValValMetLeuSerLysGlyLysThrSerGlySerThrGlyThrS 86
    |:::|||||||||:::::      ||||||:::::::::   ||||||| 
213 TGTGGTCGTCATGGTGACC......AAGACCAAAGCCGGCCAGGGTACCT 256

 86 erSerSerGlnHisSerAsnThrProAlaThrArgGlnAla..... 99
    ||:::      :::::::::      |||:::   :::::  
257 CAGCACCCCCAGAGGCCTCACCCACAGCTGCCCCAGAGTCCTCTACATCC 306

100 .................................ProProLeuGluAl 104
                                     ||||||  
307 TTCCCGCCTGCCCCCACCTCAGGCATGTCCCATCCCCCACCTGCCGCCAG 356

104 aProGlnGlnAlaProGlnProProValAlaProIleThrThrSerGlnP 121
      :::::::::|||:::         ||||||      |||::::::| 
357 AGAGGACAAGAGCCCATCAGAGGAATCCGCCCCC......ACGACGTCCC 400

121 roGluGlyLeuProAlaGlnAlaProAsnThr.............His 132
    |||||::::::   :::::::::|||::::::           ::: 
401 CAGAGTCTGTGTCAGGCTCTGTTCCCTCTTCAGGTAGCAGCGGGCGAGAG 450

133 AspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnValAspThrIleIl 149
    ::::::||||||||:::|||::::::|||   :::   :::|||:::::  
451 GAAGACGCGGCCTCCACGCTAGTGACGGGCTCTGAGTATGAGACGATGCT 500

149 eAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLysAspLysValG 166
    ::::::::::|||:::||||||         :::::::::::::|||  
501 GACGGAGATCATGTCCATGGGC.........TATGAGCGAGAGCGGGTCG 541

166 lnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluTyr 182
      ||||||||||||:::|||||||||||||:::|||||||||||||||  
542 TGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGAGCCGTGGAGTAT 591

183 LeuTyrSerGlyIleProValThrAlaGluIleAlaValProIleGlyGl 199
    |||   :::||||||||||  
592 CTGCTCACGGGAATTCCT.............................. 609

199 yGlnGlyAlaAsnThrThrAspArgAlaProThrGlyGluAlaGlyLeuS 216
      |||:::         :::::::::   :::   |||:::   :::|  
610 ....GGGAGCCCCGAGCCGGAACACGGTTCTGTCCAGGAGAGCCAGGTAT 655

216 erGlyIleProAsnThrAlaProLeuAspLeuPheProGlnGlyAlaSer 232
    ||      |||   |||  
656 CGGAGCAGCCGGCCACGGAA........................... 675

233 AsnAlaGlyGlyGlyAlaGlyGlyGlyProLeuAspPheLeuArgAsnAs 249
              :::||||||   :::||||||:::||||||||||:::::  
676 ............GCAGCAGGAGAGAACCCCCTGGAGTTCCTGCGGGACCA 713

249 nProGlnPheGlnAlaValArgGluMetValHisThrAsnProGlnIleL 266
    :||||||||||||   :::|||:::::::::::   ||||||   :::|  
714 GCCCCAGTTCCAGAACATGCGGCAGGTGATTCAGCAGAACCCTGCGCTGC 763

266 euGlnProMetLeuValGluLeuSerLysGlnAsnProGlnIleLeuArg 282
    ||      :::|||   :::|||:::::::::|||||||||:::|||:::  
764 TGCCCGCCCTGCTCCAGCAGCTGGGCCAGGAGAACCCTCAGCTTTTACAG 813

283 LeuIleGluGluAsnHisAspGluPheLeuGlnLeuLeuAsnGluPro.. 298
       |||:::::::::::::::::::|||:::|||:::||||||||||||  
814 CAAATCAGCCGGCACCAGGAGCAGTTCATCCAGATGCTGAACGAGCCCCC 863

299 ..............PheGluGlyGlyGluGlyAspPheLeuAspGlnP 310
                  ::::::      ||||||:::  
864 TGGGGAGCTGGCGGACATCTCAGATGTGGAGGGGGAGGTGGGCGCCATAG 913

310 roGluGluAspGluMetProHisAlaIleSerValThrProGluGluGln 326
```

```
      ||||||      :::   :::   |||:::||||||||||:::|||:::  
914 GAGAGGAGGCCCCGCAGATGAACTACATCCAGGTGACGCCGCAGGAGAAA 963

327 GluAlaIleGlyArgLeuGluSerMetGlyPheAspArgAlaArgValIl 343
    ||||||||   ||||||:::::::::::||||||   ::::::   ||||| 
964 GAAGCTATAGAGAGGTTGAAGGCCCTGGGCTTCCCAGAGAGCCTGGTCAT 1013

343 eGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAlaAlaAsnTyrL 360
    |:::|||:::::||||||:::::::||||||:::|||||||||||||:::|  
1014 CCAGGCCTATTTCGCGTGTGAAAAAAATGAGAACTTGGCTGCCAACTTCC 1063

360 euLeuGluHisAlaGlyGluGluAsp 368
    |||||:::::::      :::::::::  
1064 TCCTGAGTCAGAACTTTGATGACGAG 1089
```

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAF21744

seq_documentation_block:

ID AAF21744 standard; DNA; 1786 BP.
XX
AC AAF21744;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human breast and ovarian cancer associated antigen gene SEQ ID 131.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW nootropic; neurpprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS Homo sapiens.
XX
PN WO200055173-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05881.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-611515/58.
DR P-PSDB; AAB58841.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 1; Page 579; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antiulcer; vulnerary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and agonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,

```
CC   allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC   diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC   arthritis and ulcerative colitis; cardiovascular disorders such as
CC   myocardial ischaemias; wound healing; neurological diseases such as
CC   cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ   Sequence 1786 BP; 453 A; 541 C; 476 G; 316 T; 0 other;


alignment_scores:
           Quality:  603.00            Length:    392
             Ratio:  2.275               Gaps:      8
 Percent Similarity:  67.602   Percent Identity: 33.418

alignment_block:
 US-09-805-550-4 x AAF21744   ..

 Align seg 1/1  to: AAF21744  from: 1  to: 1786

       3 LeuThrValLysThrLeuLysGlyThrHisPheGluIleArgValGlnPr 19
         :::|||:::|||||||||:::         |||:::||||||:::::||
      92 ATCACGCTCAAAACGCTGCAGCAGCAGACCTTCAAGATCCGCATGGAGCC 141

      19 oAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleGlnGlyL 36
         |:::::::|||:::   :::::::|||:::::::||||||      :::|||:
     142 TGACGAGACGGTGAAGGTGCTAAAGGAGAAGATAGAAGCTGAGAAGGGTC 191

      36 ysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGlyLysVal 52
         ::|||::::::|||   :::   |||   |||||||:::   ||||||:::
     192 GTGATGCCTTCCCCGTGGCTGGACAGAAACTCATCTATGCCGGCAAGATC 241

      53 LeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAspGlyPh 69
         |||:::|||:::      :::::::::   :::::::::|||   :::||
     242 TTGAGTGACGATGTCCCTATCAGGGACTATCGCATCGATGAGAAGAACTT 291

      69 eLeuValValMetLeuSerLysGlyLysThrSerGlySerThrGlyThrS 86
         |:::|||||||||:::::      ||||||:::::::::   |||||||
     292 TGTGGTCGTCATGGTGACC......AAGACCAAAGCCGGCCAGGGTACCT 335

      86 erSerSerGlnHisSerAsnThrProAlaThrArgGlnAla..... 99
         ||:::      :::::::::      |||:::   ::::::
     336 CAGCACCCCCAGAGGCCTCACCCACAGCTGCCCCAGAGTCCTCTACATCC 385

     100 ................................ProProLeuGluAl 104
                                          ||||||
     386 TTCCCGCCTGCCCCCACCTCAGGCATGTCCCATCCCCCACCTGCCGCCAG 435

     104 aProGlnGlnAlaProGlnProProValAlaProIleThrThrSerGlnP 121
             :::::::::|||:::         ||||||      |||::::::|
     436 AGAGGACAAGAGCCCATCAGAGGAATCCGCCCCC......ACGACGTCCC 479

     121 roGluGlyLeuProAlaGlnAlaProAsnThr..............His 132
         |||||::::::   :::::::::|||::::::              :::
     480 CAGAGTCTGTGTCAGGCTCTGTTCCCTCTTCAGGTAGCAGCGGGCGAGAG 529

     133 AspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnValAspThrIleIl 149
         ::::::||||||||||:::|||:::::::|||   :::   :::|||:::::
     530 GAAGACGCGGCCTCCACGCTAGTGACGGGCTCTGAGTATGAGACGATGCT 579

     149 eAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLysAspLysValG 166
         ::::::::::|||:::||||||          ::::::::::::::|||
     580 GACGGAGATCATGTCCATGGGC.........TATGAGCGAGAGCGGGTCG 620

     166 lnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluTyr 182
             ||||||||||||:::|||||||||||:::|||||||||||||||
     621 TGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGAGCCGTGGAGTAT 670

     183 LeuTyrSerGlyIleProValThrAlaGluIleAlaValProIleGlyGl 199
         |||   :::||||||||||
     671 CTGCTCACGGGAATTCCT............................... 688

     199 yGlnGlyAlaAsnThrThrAspArgAlaProThrGlyGluAlaGlyLeuS 216
             |||:::         :::::::::   :::   |||:::   :::|

     689 ....GGGAGCCCCGAGCCGGAACACGGTTCTGTCCAGGAGAGCCAGGTAT 734

     216 erGlyIleProAsnThrAlaProLeuAspLeuPheProGlnGlyAlaSer 232
         ||       |||   |||
     735 CGGAGCAGCCGGCCACGGAA............................. 754

     233 AsnAlaGlyGlyGlyAlaGlyGlyGlyProLeuAspPheLeuArgAsnAs 249
                         ||||||   :::||||||:::|||||||||:::::
     755 ..............GCAGGAGAGAACCCCCTGGAGTTCCTGCGGGACCA 789

     249 nProGlnPheGlnAlaValArgGluMetValHisThrAsnProGlnIleL 266
         :||||||||||||   :::|||:::::::::::   ||||||   :::|
     790 GCCCCAGTTCCAGAACATGCGGCAGGTGATTCAGCAGAACCCTGCGCTGC 839

     266 euGlnProMetLeuValGluLeuSerLysGlnAsnProGlnIleLeuArg 282
         ||       :::|||   :::|||:::::::::|||||||||:::|||:::
     840 TGCCCGCCCTGCTCCAGCAGCTGGGCCAGGAGAACCCTCAGCTTTTACAG 889

     283 LeuIleGluGluAsnHisAspGluPheLeuGlnLeuLeuAsnGluPro.. 298
           |||:::::::::::::::::::|||:::|||:::||||||||||||
     890 CAAATCAGCCGGCACCAGGAGCAGTTCATCCAGATGCTGAACGAGCCCCC 939

     299 ..............PheGluGlyGlyGluGlyAspPheLeuAspGlnP 310
                         ::::::      ||||||::
     940 TGGGGAGCTGGCGGACATCTCAGATGTGGAGGGGGAGGTGGGCGCCATAG 989

     310 roGluGluAspGluMetProHisAlaIleSerValThrProGluGluGln 326
           ||||||      :::   :::   |||:::|||||||||:::|||:::
     990 GAGAGGAGGCCCCGCAGATGAACTACATCCAGGTGACGCCGCAGGAGAAA 1039

     327 GluAlaIleGlyArgLeuGluSerMetGlyPheAspArgAlaArgValIl 343
         |||||||||   ||||||:::::::::||||||   ::::::   |||||
    1040 GAAGCTATAGAGAGGTTGAAGGCCCTGGGCTTCCCAGAGAGCCTGGTCAT 1089

     343 eGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAlaAlaAsnTyrL 360
         |:::|||::::::|||||::::::||||||:::||||||||||||||:::|
    1090 CCAGGCCTATTTCGCGTGTGAAAAAAATGAGAACTTGGCTGCCAACTTCC 1139

     360 euLeuGluHisAlaGlyGluGluAsp 368
         |||||::::::      :::::::::
    1140 TCCTGAGTCAGAACTTTGATGACGAG 1165

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:AAT77781

seq_documentation_block:
ID   AAT77781 standard; cDNA; 1041 BP.
XX
AC   AAT77781;
XX
DT   01-OCT-1997  (first entry)
XX
DE   Nuclear mitotic apparatus interacting protein, NIP-2.
XX
KW   NIP-1; NIP-2; NuMA; nuclear mitotic apparatus; NuMA interacting protein;
KW   cell division; proliferation; antibody; Ab; detection;
KW   malignant cell growth; ss.
XX
OS   Homo sapiens.
XX
PN   WO9640917-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US09504.
XX
PR   07-JUN-1995;   95US-0478408.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   McPherson SMG,  Snyder MP;
XX
DR   WPI; 1997-077270/07.
DR   P-PSDB; AAW21730.
```

```
XX
PT   New nucleic acid encoding nuclear mitotic appts. interacting
PT   proteins - useful for modulating cell division and proliferation and
PT   in diagnosis
XX
PS   Claim 2; Page 64-65; 78pp; English.
XX
CC   The sequences given in AAT77780-81 encode NIP-1 and NIP-2 (NIP = NuMA
CC   (nuclear mitotic apparatus) interacting protein).  Compounds which
CC   interfere with the interaction of NuMA with a known NIP are used to
CC   modulate cell division and/or proliferation. Ab, raised conventionally
CC   using NIP-1 or -2 as immunogen, are used to detect NIP (or their
CC   complexes) and to block their activity for diagnostic or therapeutic
CC   use, e.g. to detect defective NuMA or NIP which may be markers for
CC   aberrant (including malignant) cell growth (which can also be detected
CC   by nucleic acid sequencing). Also where malignancy is related to defects
CC   in NuMA or NIP, it can be treated by admin. of the appropriate
CC   functional protein.
XX
SQ   Sequence 1041 BP; 245 A; 311 C; 314 G; 171 T; 0 other;


alignment_scores:
            Quality:  574.50              Length:    379
              Ratio:  2.244                 Gaps:      8
 Percent Similarity:  67.546    Percent Identity: 32.982

alignment_block:
 US-09-805-550-4 x AAT77781   ..

 Align seq 1/1  to: AAT77781  from: 1  to: 1041

      16 ArgValGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGl 32
         |||:::::::|||::::::|||:::   :::::::|||::::::|||||
          1 CGCATGGAGCCTGACGAGACGGTGAAGGTGCTAAAGGAGAAGATAGAAGC 50

      32 uIleGlnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheA 49
             :::|||:::|||::::::|||   :::   |||   |||||||:::
      51 TGAGAAGGGTCGTGATGCCTTCCCCGTGGCTGGACAGAAACTCATCTATG 100

      49 snGlyLysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsn 65
           ||||||:::|||:::|||:::     :::::::::   :::::::::
     101 CCGGCAAGATCTTGAGTGACGATGTCCCTATCAGGGACTATCGCATCGAT 150

      66 GluAspGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySe 82
         |||   :::|||:::||||||||||:::::      ||||||:::::::
     151 GAGAAGAACTTTGTGGTCGTCATGGTGACC......AAGACCAAAGCCGG 194

      82 rThrGlyThrSerSerSerGlnHisSerAsnThrProAlaThrArgGlnA 99
         :   ||||||||||:::      :::::::::      |||:::   ::::
     195 CCAGGGTACCTCAGCACCCCCAGAGGCCTCACCCACAGCTGCCCCAGAGT 244

      99 la.....................................Pro 100
         ::                                     |||
     245 CCTCTACATCCTTCCCGCCTGCCCCCACCTCAGGCATGTCCCATCCCCCA 294

     101 ProLeuGluAlaProGlnGlnAlaProGlnProProValAlaProIleTh 117
         |||          ::::::::::|||:::          ||||||
     295 CCTGCCGCCAGAGAGGACAAGAGCCCATCAGAGGAATCCGCCCCC..... 339

     117 rThrSerGlnProGluGlyLeuProAlaGlnAlaProAsnThr..... 131
          |||::::::||||||::::::   ::::::::::|||::::::
     340 .ACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCTTCAGGTAGCA 388

     132 ........HisAspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnVal 145
                 :::::::::|||||||||:::|||::::::|||   :::
     389 GCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGGCTCTGAGTAT 438

     146 AspThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLy 162
         :::|||:::::::::::::::|||:::||||||         ::::::::
     439 GAGACGATGCTGACGGAGATCATGTCCATGGGC.........TATGAGCG 479

     162 sAspLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgA 179
         :::::::|||      ||||||||||||:::||||||||||||:::||||
     480 AGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCGAG 529

     179 laValGluTyrLeuTyrSerGlyIleProValThrAlaGluIleAlaVal 195
         ||||||||||||||   :::|||||||||
     530 CCGTGGAGTATCTGCTCACGGGAATTCCT..................... 558

     196 ProIleGlyGlyGlnGlyAlaAsnThrThrAspArgAlaProThrGlyGl 212
                     |||:::         :::::::::   :::   ||
     559 ..............GGGAGCCCCGAGCCGGAACACGGTTCTGTCCAGGA 593

     212 uAlaGlyLeuSerGlyIleProAsnThrAlaProLeuAspLeuPheProG 229
         |:::   :::|||       |||   |||
     594 GAGCCAGGTATCGGAGCAGCCGGCCACGGAA................... 624

     229 lnGlyAlaSerAsnAlaGlyGlyGlyAlaGlyGlyGlyProLeuAspPhe 245
                                 :::|||||||   :::||||||:::|||
     625 ......................GCAGCAGGAGAGAACCCCCTGGAGTTC 651

     246 LeuArgAsnAsnProGlnPheGlnAlaValArgGluMetValHisThrAs 262
         ||||||:::::|||||||||||||   :::|||||||||||||   ||
     652 CTGCGGGACCAGCCCCAGTTCCAGAACATGCGGCAGGTGATTCAGCAGAA 701

     262 nProGlnIleLeuGlnProMetLeuValGluLeuSerLysGlnAsnProG 279
         ||||   :::|||       :::|||   :::|||:::::::::|||||||
     702 CCCTGCGCTGCTGCCCGCCCTGCTCCAGCAGCTGGGCCAGGAGAACCCTC 751

     279 lnIleLeuArgLeuIleGluGluAsnHisAspGluPheLeuGlnLeuLeu 295
         ||:::|||:::   |||:::::::::::::::::::|||:::|||:::|||
     752 AGCTTTTACAGCAAATCAGCCGGCACCAGGAGCAGTTCATCCAGATGCTG 801

     296 AsnGluPro.................PheGluGlyGlyGluGlyAspPh 306
         |||||||||                 ::::::      |||||||:::
     802 AACGAGCCCCCTGGGGAGCTGGCGGACATCTCAGATGTGGAGGGGGAGGT 851

     306 eLeuAspGlnProGluGluAspGluMetProHisAlaIleSerValThrP 323
                       ||||||      :::   :::   |||:::|||||||
     852 GGGCGCCATAGGAGAGGAGGCCCCGCAGATGAACTACATCCAGGTGACGC 901

     323 roGluGluGlnGluAlaIleGlyArgLeuGluSerMetGlyPheAspArg 339
         ||:::|||:::|||||||||   ||||||:::::::::||||||   :::
     902 CGCAGGAGAAAGAAGCTATAGAGAGGTTGAAGGCCCTGGGCTTCCCAGAG 951

     340 AlaArgValIleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAl 356
         :::   ||||||:::|||:::::::||||||::::::||||||:::||||
     952 AGCCTGGTCATCCAGGCCTATTTCGCGTGTGAAAAAAATGAGAACTTGGC 1001

     356 aAlaAsnTyrLeuLeuGluHisAlaGlyGluGluAsp 368
         |||||||:::||||||::::::     :::::::::
    1002 TGCCAACTTCCTCCTGAGTCAGAACTTTGATGACGAG 1038

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:AAC44960

seq_documentation_block:
ID   AAC44960 standard; DNA; 850 BP.
XX
AC   AAC44960;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 44778.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
```

XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

```
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
           Quality:  563.50              Length:    258
             Ratio:  2.997                 Gaps:      5
 Percent Similarity:  72.868   Percent Identity: 48.062

alignment_block:
 US-09-805-550-4 x AAC44960   ..

 Align seg 1/1  to: AAC44960  from: 1  to: 850

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         |||||||:::   |||||||||||||:::|||:::::::||||||||:::||
      79 ATGAAGATTTTCGTGAAGACTCTCAGTGGTTCGAACTTCGAGATCGAAGT 128

      17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
         |:::|||   |||   :::      ||||||      ||||||   ::::
     129 AAAACCTGCAGATAGGGTTTCTGATGTTAAAACGGCTATAGAAACTGTTA 178

      34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
         ::|||   :::   ||||||   :::::::|||:::||||||   :::|||
     179 AAGGTGCAGAA...TACCCTGCTGCTAAGCAGATGCTGATCCACCAAGGA 225

      51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
         ||||||||||||||||||:::|||||||||||||||||:::|||   |||::
     226 AAGGTTCTAAAGGATGAGACCACATTGGAAGAGAACAATGTTGTTGAGAA 275

      67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSer.........G 81
         :::::|||:::|||:::||||||||||||   |||:::|||         :
     276 CAGTTTCATTGTTATCATGTTGTCCAAGACCAAGGCTTCTNCAAGTGGGG 325

      81 lySerThrGlyThrSerSerSerGlnHisSerAsnThrProAlaThrArg 97
         ::||||||:::::::::   :::      ::::::   |||  .|||
     326 CATCAACCGCATCTGCACCAGCACCTAGTGCTACACAGCCACAGACAGTA 375

      98 GlnAlaProProLeuGluAlaPro...GlnGlnAlaProGlnProProVa 113
         :::|||   ::::::||||||      :::::::|||   |||
     376 GCTACACCTCAGGTTTCTGCTCCTACTGCCTCAGTTCCAGTTCCTACAAG 425

     113 lAlaProIleThrThrSerGlnProGluGlyLeuProAlaGlnAlaProA 130
          :::      |||::::::   |||           :::        :
     426 TGGAACTGCAACCGCTGCAGCTCCAGCTACTGCTGCTTCGGTTCAGACAG 475

     130 snThrHisAspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnValAsp 146
         ::::::::   :::||||||||||||||:::::|||   :::::::::
     476 ATGTGTATGGTCAAGCAGCATCAAACCTTGTTGCTGGAACTACTTTAGAG 525

     147 ThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySerTrpAspLysAs 163
         :::   ::::::|||:::::::::|||||||||||||||||||:::||
     526 TCCACTGTTCAGCAAATTCTTGATATGGGTGGAGGTAGTTGGGACCGTGA 575

     163 pLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnProGluArgAlaV 180
         |   |||   ||||||||||||||||||:::|||||||||||||||||||
     576 CACTGTTGTTCGTGCTCTGAGAGCTGCGTTCAACAATCCTGAAAGAGCTG 625

     180 alGluTyrLeuTyrSerGlyIleProValThrAlaGluIle......... 193
         ||||||||||||||||||||||||:::   |||||||||
     626 TTGAATACCTGTACTCGGGTATCCCTGCTCAAGCTGAAATCCCACCAGTT 675

     194 ......AlaValProIleGlyGlyGlnGlyAlaAsnThrThrAspArgAl 208
               |||         |||   |||:::||||||      :::
     676 GCTCAAGCCCCAGCTACTGGGGAACAGGCAGCCAATCCTCTAGCACAG.. 723

     208 aProThrGlyGluAlaGlyLeuSerGlyIleProAsnThrAlaProLeuA 225
          |||         |||:::   :::|||   |||||||:::   ||||||:
     724 .CCCCAACAAGCAGCAGCTGCAACTGGTGGTCCAAACGCAAATCCATTAA 772

     225 spLeuPheProGlnGlyAlaSerAsnAlaGlyGlyGlyAlaGlyGlyGly 241
         ::|||||||||||||||         |||   :::|||||||||:::|||
     773 ACCTGTTCCCCCAGGGCATGCCCGCTGCAGATGCTGGTGCTGGAGCTGGT 822

     242 ProLeuAspPheLeuArgAsnAsn 249
             ||||||||||||||||||:::
     823 AATCTTGATTTCCTGCGTAACAGT 846

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABL03535

seq_documentation_block:
ID   ABL03535 standard; cDNA; 1477 BP.
XX
AC   ABL03535;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 5087.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB59432.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
```

```
PS   Claim 1; SEQ ID NO 5087; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1477 BP; 462 A; 313 C; 298 G; 404 T; 0 other;


alignment_scores:
            Quality:  487.50              Length:    445
              Ratio:  1.935                Gaps:     11
 Percent Similarity:  56.629   Percent Identity: 28.764

alignment_block:
 US-09-805-550-4 x ABL03535   ..

 Align seg 1/1  to: ABL03535  from: 1  to: 1477

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         |||   :::|||:::|||:::|||:::          |||   |||:::
     118 ATGATTATTACAATTAAAAATCTTCAACAGCAAACTTTTACTATTGAGTT 167

      17 lGlnProAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleG 34
             |||:::   |||::::::   :::||||||:::|||   |||   :
     168 TGCCCCGGAAAAAACGGTTTTGGAACTGAAGAAGAAAATATTCGAGGAGC 217

      34 lnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGly 50
         ::|||   :::   |||         :::|||   ||||||:::   |||
     218 GCGGTCCAGAG...TACGTCGCCGAAAAACAAAAACTGATCTACGCTGGC 264

      51 LysValLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAs 67
            :::|||   |||:::   |||:::   :::   :::|||:::|||
     265 GTCATATTGACGGATGACCGCACCGTTGGTTCATACAACGTTGATGAAAA 314

      67 pGlyPheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrG 84
             |||:::|||||||||||||:::::   ::::::|||:::::   :
     315 AAAGTTCATTGTGGTAATGTTGACACGCGATTCGTCTAGTTCAAATCGTA 364

      84 lyThrSerSerSerGlnHisSerAsn.....ThrProAlaThr 96
         ::      |||   ::::::||||||          |||      :::
     365 ATCAACTTAGTGTAAAAGAAAGTAATAAATTGACTAGCACCGACGATTCA 414

      97 ArgGlnAlaProProLeuGluAlaProGlnGlnAlaProGlnPro..... 111
         :::|||:::   |||   |||      :::::::::   :::|||
     415 AAGCAATCTATGCCTTGCGAAGAAGCCAACCATACCAATTCGCCTAGTTC 464

     112 .............................ProValAlaProIleThrT 118
                                        :::   |||::::::::
     465 CACAAATACAGAAGATTCAGTTTTATCACGTGAAACCAGACCTTTATCTA 514

     118 hrSerGlnProGluGlyLeuProAlaGlnAlaProAsnThrHisAspAsn 134
         ::::::::      |||      |||||||||   :::   :::::::::
     515 GTGACGAATTGATTGGCGAGTTGGCCCAGGCT...TCCTTACAATCGCGC 561

     135 AlaAlaSerAsnLeuLeuSerGlyArgAsnValAspThrIleIleAsnGl 151
         |||   ||||||||||||   |||   :::   :::      :::   ::
     562 GCTGAATCTAATTTGCTTATGGGTGACGAATACAACCAAACAGTCCTATC 611

     151 nLeuMetGluMetGlyGlyGlySerTrpAspLysAspLysValGlnArgA 168
         :::::::|||||||||         :::   :::::::::|||:::||||
     612 AATGGTGGAAATGGGT.........TACCCAAGAGAGCAGGTTGAACGTG 652

     168 laLeuArgAlaAlaTyrAsnAsnProGluArgAlaValGluTyrLeuTyr 184
         ||:::   |||:::|||||||||||||||||||||||||||||||||||

     653 CGATGGCTGCTAGTTATAACAACCCGGAAAGAGCCGTTGAATATCTCATT 702

     185 SerGlyIleProVal................................... 189
         :::|||||||||:::
     703 AATGGCATACCTGCAGAGGAAGGTACTTTTTACAATAGGCTGAATGAATC 752

     190 .ThrAlaGluIleAlaValProIleGlyGlyGlnGlyAlaAsnThrThrA 206
          |||            :::|||   |||   |||   |||::::::|||:
     753 AACGAATCCTAGTCTAATCCCCTCCGGACCGCAACCTGCGTCGGCAACCT 802

     206 spArgAlaProThrGlyGluAlaGlyLeuSerGlyIleProAsnThrAla 222
         ::         :::   |||
     803 CTGCCGAGCGTTCAACAGAA.............................. 822

     223 ProLeuAspLeuPheProGlnGlyAlaSerAsnAlaGlyGlyGlyAlaGl 239
                                                     :::::
     823 ..........................................TCAAA 827

     239 yGlyGlyProLeuAspPheLeuArgAsnAsnProGlnPheGlnAlaValA 256
         ::::   |||::::::|||||||||:::::|||||||||      :::|
     828 TTCAGACCCCTTTGAATTTTTACGTAGCCAGCCACAGTTTCTTCAAATGC 877

     256 rgGluMetValHisThrAsnProGlnIleLeuGlnProMetLeuValGlu 272
         ||:::::::::::::   ||||||:::::::|||:::   :::|||   :::
     878 GATCTTTAATTTATCAAAACCCTCATCTTTTGCATGCAGTATTGCAGCAG 927

     273 LeuSerLysGlnAsnProGlnIleLeuArgLeuIleGluGluAsnHisAs 289
         :::::::::   ||||||   :::|||:::||||||:::||||||:::||
     928 ATTGGTCAGACAAACCCAGCCCTTTTACAACTTATTTCGGAGAACCAAGA 977

     289 pGluPheLeuGlnLeuLeuAsnGluProPhe.............GluGlyG 302
         |   ||||||::::::||||||:::|||:::                |||:::|
     978 TGCGTTTCTCAATATGCTTAATCAACCGATTGACCGCGAATCGGAGTCAG 1027

     302 ly.............................................. 302
         ||
    1028 GCGCCACCGTTCCTCCCGTTTCGAATGCGCGGATTCCTTCAACTTTGGAC 1077

     302 .............................................. 302

    1078 AATGTTGACCTCTTTTCTCCAGATTTAGAAGTAGCTACTTCAGCACAAAG 1127

     303 ............................GluGlyAspPheLeuAspG 309
                                       :::|||:::      |||:
    1128 ATCAGCCGCGGGTACAAGTGCAGCACATCAAAGCGGTAGCGCAGCGGATA 1177

     309 lnProGluGluAspGluMetProHisAlaIleSer.........ValThr 322
         ::   |||:::   |||:::|||   ::::::|||         ::::::
    1178 AC...GAAGACTTGGAACAACCTTTAGGAGTATCAACCATTCGTTTAAAT 1224

     323 ProGluGluGlnGluAlaIleGlyArgLeuGluSerMetGlyPheAspAr 339
         ::::::::::::::|||||   |||||:::::::::|||||   ::
    1225 CGCCAAGATAAGGACGCAATAGAACGGCTAAAAGCTCTTGGATTCCCAGA 1274

     339 gAlaArgValIleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuA 356
         :|||   |||:::::::|||:::::::||||||:::::|||||||||   |
    1275 GGCCCTTGTACTGCAAGCGTACTTCGCTTGTGAAAAGAACGAGGAACAAG 1324

     356 laAlaAsnTyrLeuLeuGluHisAlaGlyGluGlu 367
         ||||||||:::||||||:::   :::   ::::::
    1325 CAGCTAATTTTTTGCTATCGTCTAGCTTCGATGAT 1359

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:AAS78241

seq_documentation_block:
ID   AAS78241 standard; cDNA; 944 BP.
XX
AC   AAS78241;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #14045.
XX
```

```
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG14054.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 14045; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 944 BP; 204 A; 303 C; 279 G; 158 T; 0 other;


alignment_scores:
            Quality:  385.00              Length:    317
              Ratio:  1.869                 Gaps:     14
 Percent Similarity:  64.984   Percent Identity: 32.808

alignment_block:
 US-09-805-550-4 x AAS78241   ..

 Align seq 1/1  to: AAS78241  from: 1  to: 944

       3 LeuThrValLysThr.LeuLysGlyThrHisPheGluIleArgValGlnP 19
         :::||||||||||||| |||:::         |||:::||||||:::::::|
     100 ATCACGGTCAAAACCGCTGCAGCAGCAGACCTTCAAGATCCGCATGGAGC 149

      19 roAsnAspThrIleMetAlaValLysLysAsnIleGluGluIleGlnGly 35
         ||:::::::|||:::   :::::::|||:::::::||||||      :::|||
     150 CTGACGAGACGGTGAAGGTGCTAAAGGAGAAGATAGAAGCTGAGAAGGGT 199

      36 .LysAspSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGlyLysV 52
          :::|||:::::::|||   :::   |||   |||||||:::   ||||||:

     200 CCGTGATGCCTTCCCCGTGGCTGGACAGAAACTCATCTATGCCGGCAAGA 249

      52 alLeuLysAspGluSerThrLeuGluGluAsnLysValAsnGluAspGly 68
         ::|||:::|||:::      :::::::::   ::::::::::|||   :::
     250 TCTTGAGTGACGATGTCCCTATCAGGGACTATCGCATCGATGAGAAGAAC 299

      69 PheLeuValValMetLeuSerLysGlyLysThrSerGlySerThrGlyTh 85
         |||:::|||||||||:::::      ||||||:::::::::   |||||
     300 TTTGTGGTCGTCATGGTGACC......AAGACCAAAGCCGGCCAGGGTAC 343

      85 rSerSerSerGlnHisSerAsnThrProAlaThrArgGlnAlaPro.... 100
         ||||:::      ::::::   |||   :::      |||:::|||
     344 CTCAGCACCCCCAGAGGCCTTCACCCACAGCTTGCCCCAGAGTCCTCTTA 393

     101 ........Pro..LeuGluAlaProGlnGln...AlaProGlnProPro. 112
                 .|||  |||      ||||||      :::|||:::||||||
     394 CATCCTTCCCCGCCTGCCCCCACCTCAGGGCATTGTTCCCCATCCCCCAC 443

     113 ....:.................................ValAlaProIle 116
                                               ||||||
     444 CTTGCCGCCAGAGAGGGGACAAGAGCCCATCAGAGGAATCCGCCCCC... 490

     117 ThrThrSerGlnProGluGlyLeuProAlaGlnAlaProAsnThr..... 131
            |||:::::::||||||:::::   :::::::::|||:::::
     491 ...ACGACGTCCCCAGAGTCTGTGTCAGGCTCTGTTCCCTCTTCAGGTAG 537

     132 ..........HisAspAsnAlaAlaSerAsnLeuLeuSerGlyArgAsnV 145
                   :::::::::|||||||||:::|||:::::::|||   :::
     538 CAGCGGGCGAGAGGAAGACGCGGCCTCCACGCTAGTGACGGGCTCTGAGT 587

     145 alAspThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySerTrpAsp 161
           :::|||:::::::::::::::::|||:::||||||          ::::::
     588 ATGAGACGATGCTGACGGAGATCATGTCCATGGGC.........TATGAG 628

     162 LysAspLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnProGluAr 178
         :::::::::|||      ||||||||||||:::|||||||||||||:::||
     629 CGAGAGCGGGTCGTGGCCGCCCTGAGAGCCAGCTACAACAACCCCCACCG 678

     178 gAlaValGluTyrLeuTyrSerGlyIleProValThrAlaGluIleAlaV 195
         ||||||||||||||||   :::||||||||||
     679 AGCCGTGGAGTATCTGCTCACGGGAATTCCT..................... 709

     195 alProIleGlyGlyGlnGlyAlaAsnThrThrAspArgAlaProThrGly 211
                         |||:::          :::::::::
     710 .....:..........GGGAGCCCCGAGCCGGAACACGGTTCTTTCCAG 742

     212 GluAlaGlyLeuSerGlyIleProAsnThrAlaProLeuAspLeuPhePr 228
         |||:::   :::|||      |||   |||
     743 GAGAGCCAGGTATCGGAGCAGCCGGCCACGGAA.................. 775

     228 oGlnGlyAlaSerAsnAlaGlyGlyGlyAlaGlyGlyGlyProLeu.Asp 244
                                    :::|||||   :::|||||| :::
     776 .......................GCAGCAGGAGAGAACCCCCTGGGAG 800

     245 PheLeuArgAsnAsnProGlnPheGlnAlaValArgGluMetVal.HisT 261
         ||||||||:::::|||||||||||   :::|||:::::::::: :::
     801 TTCCTGCGGGACCAGCCCCAGTTCCAGAACATGCGGCAGGTGATTTCAGC 850

     261 hrAsnProGlnIleLeuGlnProMetLeuValGluLeuSerLysGlnAsn 277
           ||||||   :::|||      :::|||   :::|||:::::::::|||
     851 AGAACCCTGCGCTGCTGCCCGCCCTGCTCCAGCAGCTGGGCCAGGAGAAC 900

     278 ProGlnIleLeuArgLeuIleGluGluAsnHisAspGluPhe 291
         ||||||:::|||:::   |||:::::::::::::::::|||
     901 CCTCAGCTTTTACAGCAAATCAGCCGGCACCAGGAGCAGTTC 942

seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABL16782

seq_documentation_block:
ID   ABL16782 standard; DNA; 6037 BP.
XX
AC   ABL16782;
XX
```

```
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 1819.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 1819; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 6037 BP; 1525 A; 1355 C; 1388 G; 1769 T; 0 other;


alignment_scores:
           Quality:  375.00            Length:    376
             Ratio:  1.752               Gaps:     14
 Percent Similarity:  56.915   Percent Identity: 26.596

alignment_block:
 US-09-805-550-4 x ABL16782   ..

 Align seg 1/1  to: ABL16782  from: 1  to: 6037

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         |||||||||:|||||||||   |||            :::   ::::::::
    3540 ATGAAGCTGTCTATACGCATGCTGGACCAACGCACCATCACTTTGGAGAT 3589

      17 lGlnProAsnAspThrIleMetAlaValLysLys.........AsnIleG 31
         ::::   ::::::   :::   |||:::|||:::         |||:::
    3590 GAACGAATCGCAGGAGGTGAGGGCTCTGAAGCAGAAATTGGGCAATTTAC 3639

      31 luGluIleGlnGlyLysAspSerTyrProTrpGlyGlnGlnLeuLeuIle 47
         |||:::           :::   |||      :::      ||||||
    3640 CCGAAGTC............GCCATGCCCGCGGAGAACCTTCAGCTGATA 3677

      48 PheAsnGlyLysValLeuLysAspGluSerThrLeuGluGluAsnLysVa 64
         :::::::|||:::::::::::::|||         |||:::|||   :::::
    3678 TACAGTGGCCGCATTATGGAGGATGCCATGCCCCTCAGTGAATATCGTAT 3727

      64 lAsnGluAspGlyPheLeuValValMetLeuSerLysGlyLysThrSerG 81
         :   ||||||      :::::::|||:::::::::|||   |||::::::

    3728 AGCCGAGGACAAG...ATCATTGTGTTGATGGGTAAGAAGAAGGTTGATA 3774

      81 lySerThrGlyThrSerSerSerGlnHisSerAsnThrProAlaThrArg 97
         |||
    3775 AGAGC..................................................... 3779

      98 GlnAlaProProLeuGluAlaProGlnGln......AlaProGlnProPr 112
                           :::|||::::::      ||||||   |||||
    3780 .................TCGCCAGAGGAGAAGGTTGCCCCGACACCACC 3811

     112 oValAla......ProIleThrThrSerGlnProGluGlyLeuProAlaG 127
         |:::|||      |||   :::            :::   :::|||:::
    3812 GTTGGCCGCTGGCCCAAATGTTTTGCGCACAGAGGATGTGGTGCCTTCAC 3861

     127 lnAlaProAsnThrHisAspAsnAlaAlaSerAsnLeuLeuSerGlyArg 143
           |||||||||
    3862 TAGCTCCCAAT.................................... 3872

     144 AsnValAspThrIleIleAsnGlnLeuMetGluMetGlyGlyGlySerTr 160
            |||      :::::::::||||||:::||||||   |||
    3873 ......GATCAGTGGGTGAGCGATCTCATGTCAATGGGATATGGC..... 3911

     160 pAspLysAspLysValGlnArgAlaLeuArgAlaAlaTyrAsnAsnProG 177
             :::::::::|||:::   ||||||||||||:::::::|||:::||||
    3912 ....GAAGAGGAGGTACGCTCAGCCCTCCGGGCGAGCTTTAATCATCCGG 3957

     177 luArgAlaValGluTyrLeuTyrSerGlyIleProValThrAlaGluIle 193
         ||||||||:::|||||||||   :::|||||||||
    3958 AAAGGGCTATAGAGTATTTGATTAATGGGATTCCT.............. 3992

     194 AlaValProIleGlyGlyGlnGlyAlaAsnThrThrAspArgAlaProTh 210
                                                 :::      ::
    3993 .....................................CAGGAGGTTGT 4003

     210 rGlyGluAlaGlyLeuSerGlyIleProAsnThrAlaProLeuAspLeuP 227
         :::|||   ||||||:::::||||||:::::          |||   :
    4004 TTCAGAGCAGGGATTAGCTGCAATCCCGAGCGTACAGACAAGTGATCAAT 4053

     227 heProGlnGlyAlaSerAsnAlaGlyGlyGlyAlaGlyGlyGlyProLeu 243
         ::   |||
    4054 TGCAGCAA..................................... 4061

     244 AspPheLeuArgAsnAsnProGlnPheGlnAlaValArgGluMetValHi 260
            |||      :::   ::::::      :::||||||||||:::::
    4062 ......TTAATGGCAGATCTTAACATTACACGGATGCGTGAGATGATTAA 4105

     260 sThrAsnProGlnIleLeuGlnProMetLeuValGluLeuSerLysGlnA 277
         :   ||||||:::::::::::::   ::::::   :::|||:::::   :
    4106 TCAGAATCCAGAACTAATACACAGACTAATGAACAGACTGGCTGAAACCG 4155

     277 snProGlnIleLeuArgLeuIleGluGluAsnHisAspGluPheLeuGln 293
         ::|||      :::::::::::::::::|||::::::|||:::::::::
    4156 ATCCGGCTACCTTCGAAGTCTTTCAGCGTAACCAGGAGGAGTTAATGAAC 4205

     294 LeuLeuAsnGluProPheGluGlyGlyGluGlyAspPheLeuAspGlnPr 310
         :::::::::            ||||||                  :::
    4206 ATGATTTCA...........GGCGGCGCA.............AGTCG 4228

     310 oGluGluAspGluMetProHis...AlaIleSerValThrProGluGluG 326
                ::::|||:::   |||      |||::::::|||   ||||||
    4229 CACCCCGAACGAGATTGAACATTTACAGATTACTTTAACCGCCGAAGAAA 4278

     326 lnGluAlaIleGlyArgLeuGluSerMetGlyPheAspArgAlaArgVal 342
           |||:::||||||||||||:::::::||||||:::|||:::   :::
    4279 CCGCCGCCGTAGGGCGTTTGGAGGCACTGGGTTTCGAACGTGTGATGGCC 4328

     343 IleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAlaAlaAsnTy 359
         ::::::|||:::||||||||||||:::::|||:::||||||||||:::
    4329 GTTCAGGCCTATCTGGCCTGCGACAAGGACGAGCAGCTGGCCGCAGAGGT 4378

     359 rLeuLeuGluHisAlaGlyGluGluAsp 368
           |||:::::::::::   |||||||||
    4379 ACTAATACGCCAGTCA...GAAGAGGAT 4403
```

```
seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:ABL03534

seq_documentation_block:
ID   ABL03534 standard; cDNA; 4198 BP.
XX
AC   ABL03534;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 5084.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB59431.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 5084; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 4198 BP; 1397 A; 710 C; 719 G; 1372 T; 0 other;


alignment_scores:
            Quality:  346.50              Length:    661
              Ratio:  1.274                 Gaps:     19
 Percent Similarity:  41.150   Percent Identity: 20.575

alignment_block:
 US-09-805-550-4 x ABL03534/rev   ..

 Align seg 1/1  to reverse of: ABL03534  from: 1  to: 4198

       1 MetLysLeuThrValLysThrLeuLysGlyThrHisPheGluIleArgVa 17
         |||   :::|||:::|||:::|||:::          |||   |||:::
    3081 ATGATTATTACAATTAAAAATCTTCAACAGCAAACTTTTACTATTGAGTT 3032

      17 lGlnProAsnAspThr.............................. 22
             |||:::   |||
    3031 TGCCCCGGAAAAAACGGTATGTGTACAAATAGTCAAATACTAGTATATTC 2982

      22 ..................................................... 22
```

```
2981 ACATGTATGGTATATACAATTATTGCAAATGCATATGTGTGCTTATACAG 2932

  22 ..................................................... 22

2931 GTATTTCTTACAATTAATTTGTTGTACCATAATAAGCAGGGATTCTGATT 2882

  23 .......IleMetAlaValLysLysAsnIleGluGluIleGlnGlyLysA 37
           ::::::   :::||||||:::|||   |||   :::|||   :
2881 ACAAAAGGTTTTGGAACTGAAGAAGAAAATATTCGAGGAGCGCGGTCCAG 2832

  37 spSerTyrProTrpGlyGlnGlnLeuLeuIlePheAsnGlyLysValLeu 53
     ::   |||         :::|||   ||||||:::   |||   :::|||
2831 AG...TACGTCGCCGAAAAACAAAAACTGATCTACGCTGGCGTCATATTG 2785

  54 LysAspGluSerThrLeuGluGluAsnLysValAsnGluAspGlyPheLe 70
        |||:::   |||:::   :::   :::|||:::|||       |||::
2784 ACGGATGACCGCACCGTTGGTTCATACAACGTTGATGAAAAAAAGTTCAT 2735

  70 uValValMetLeuSerLysGlyLysThrSerGlySerThrGlyThrSerS 87
     :||||||||||||::::::   ::::::|||::::::   :::      |
2734 TGTGGTAATGTTGACACGCGATTCGTCTAGTTCAAATCGTAATCAACTTA 2685

  87 erSerGlnHisSerAsn............ThrProAlaThrArgGlnAla 99
     ||   ::::::||||||            |||      ::::::|||:::
2684 GTGTAAAAGAAAGTAATAAATTGACTAGCACCGACGATTCAAAGCAATCT 2635

 100 ProProLeuGluAlaProGlnGlnAlaProGlnPro.............. 111
        |||   |||      :::::::::   :::|||
2634 ATGCCTTGCGAAGAAGCCAACCATACCAATTCGCCTAGTTCCACAAATAC 2585

 112 ......................ProValAlaProIleThrThrSerGlnP 121
                                  :::   |||:::::::::::::::
2584 AGAAGATTCAGTTTTATCACGTGAAACCAGACCTTTATCTAGTGACGAAT 2535

 121 roGluGlyLeuProAlaGlnAlaProAsnThrHisAspAsnAlaAlaSer 137
         |||       |||||||||   :::   ::::::::::|||   |||
2534 TGATTGGCGAGTTGGCCCAGGCT...TCCTTACAATCGCGCGCTGAATCT 2488

 138 AsnLeuLeuSerGlyArgAsnValAspThrIleIleAsnGlnLeuMetGl 154
     |||||||||   |||   :::   :::       :::   ::::::::::||
2487 AATTTGCTTATGGGTGACGAATACAACCAAACAGTCCTATCAATGGTGGA 2438

 154 uMetGlyGlyGlySerTrpAspLysAspLysValGlnArgAlaLeuArgA 171
     |||||||         :::   ::::::::::|||:::||||||:::   |
2437 AATGGGT.........TACCCAAGAGAGCAGGTTGAACGTGCGATGGCTG 2397

 171 laAlaTyrAsnAsnProGluArgAlaValGluTyrLeuTyrSerGlyIle 187
     ||:::||||||||||||||||||||||||||||||||   :::||||||
2396 CTAGTTATAACAACCCGGAAAGAGCCGTTGAATATCTCATTAATGGCATA 2347

 188 ProVal.....................................ThrAlaGl 192
     |||:::                                    |||
2346 CCTGCAGAGGAAGGTACTTTTTACAATAGGCTGAATGAATCAACGAATCC 2297

 192 uIleAlaValProIleGlyGlyGlnGlyAlaAsnThrThr......AspA 207
           :::|||   |||   |||   |||:::::::|||      :::|
2296 TAGTCTAATCCCCTCCGGACCGCAACCTGCGTCGGCAACCTCTGCCGAGC 2247

 207 rgAlaProThrGlyGluAlaGly...LeuSerGlyIleProAsnThrAla 222
     ||:::      :::::::::|||   |||   |||:::   :::::::::
2246 GTTCAACAGAATCAAATTCAGGTAACTTGATGGGATTTCAATCAGCAGTC 2197

 223 ProLeuAspLeu.........................PheProGlnGl 230
        |||   :::                          :::|||:::::
2196 AATTTATGTATTCTTGTATGAAACGAAAATATATAATTTCTACCCAGAAG 2147

 230 yAlaSerAsnAlaGly...............................G 236
     :::::::   |||                                  |
2146 TACTAAATTTGCACACTTCAAAACATAGTGAAATTAATAAACCCGATTTG 2097

 236 lyGlyAlaGly..................................... 239
     ||:::   :::
```

```
2096 GCTCTCGCTCTTTAATGAAATTATACGGCCTTAACATAATTGCGTTAGTT 2047

 239 ..................................................  239

2046 TTCCTGGTCAAGAAATATTGTTGATCAAAACTTATGTTGCACATTCGTTA 1997

 239 ..................................................  239

1996 ATAAATAATAAAAGAGTTCTCTTAGTACTGGAAATATATGCAGCTGTTAA 1947

 239 ..................................................  239

1946 AGATTCACTAAAATGTTTGTGAGCTCGTTGTCCCGCAGTATCATTTACAT 1897

 239 ..................................................  239

1896 ATATATGTCAACAGACACTAAAGGAGCGATAATTCACGAATTTTTGTCAC 1847

 239 ..................................................  239

1846 TATTCTCTTCAAGGACAGGGCATCACTATATGTTTTTAGTTAATTTAATT 1797

 240 ........................GlyGlyProLeuAspPheLeuArg 247
                              :::   |||::::::|||||||||
1796 ATTATTTAATTATTTATATTTACGTTTCAGACCCCTTTGAATTTTTACGT 1747

 248 AsnAsnProGlnPheGlnAlaValArgGluMetValHisThrAsnProGl 264
     ::::::|||||||||      :::|||:::::::::::   ||||||::
1746 AGCCAGCCACAGTTTCTTCAAATGCGATCTTTAATTTATCAAAACCCTCA 1697

 264 nIleLeuGlnProMetLeuValGluLeuSer..................... 274
     ::::|||:::   :::|||   :::::::|||
1696 TCTTTTGCATGCAGTATTGCAGCAGGTAAGTACACTTATAATAATAATTA 1647

 275 .....................................LysGlnAsn 277
                                          :::   |||
1646 CATGACACGTAAATGATAATTTTTTTTACGTATAGATTGGTCAGACAAAC 1597

 278 ProGlnIleLeuArgLeuIleGluGluAsnHisAspGluPheLeuGlnLe 294
     |||   :::|||:::||||||:::||||||:::|||   ||||||:::::
1596 CCAGCCCTTTTACAACTTATTTCGGAGAACCAAGATGCGTTTCTCAATAT 1547

 294 uLeuAsnGluProPhe............GluGlyGly.............. 302
     :||||||:::|||:::            |||:::|||
1546 GCTTAATCAACCGATTGACCGCGAATCGGAGTCAGGCGCCACCGTTCCTC 1497

 302 ..................................................  302

1496 CCGTTTCGAATGCGCGGATTCCTTCAACTTTGGACAATGTTGACCTCTTT 1447

 302 ..................................................  302

1446 TCTCCAGATTTAGAAGTAGCTACTTCAGCACAAAGATCAGCCGCGGGTAC 1397

 303 ...............GluGlyAspPheLeuAspGlnProGluGluAspG 314
                   :::|||:::      |||:::   |||:::   |
1396 AAGTGCAGCACATCAAAGCGGTAGCGCAGCGGATAAC...GAAGACTTGG 1350

 314 luMetProHisAlaIleSer.........ValThrProGluGluGlnGlu 327
     ||:::|||   :::::::|||         ::::::   ::::::::::::
1349 AACAACCTTTAGGAGTATCAACCATTCGTTTAAATCGCCAAGATAAGGAC 1300

 328 AlaIleGlyArgLeu................................... 332
     ||||||   |||:::
1299 GCAATAGAACGGGTATATAGGAGTGAATAATTATATATGGACTGTCGAGT 1250

 333 .........................GluSer.MetGlyPheAspArgAla 340
                              :::||| :::||||||   :::|||
1249 TACTTAATATATTCTTCTTCTAGCTAAAAGCTCTTGGATTCCCAGAGGCC 1200

 341 ArgValIleGluAlaPheLeuAlaCysAspArgAsnGluGluLeuAlaAl 357
        |||:::::::|||::::::|||||||:::::|||||||||   |||||
1199 CTTGTACTGCAAGCGTACTTCGCTTGTGAAAAGAACGAGGAACAAGCAGC 1150

 357 aAsnTyrLeuLeuGluHisAlaGlyGluGlu 367
     ||||:::|||||||:::   :::   ::::::
1149 TAATTTTTTGCTATCGTCTAGCTTCGATGAT 1119
```